

176528

STIC-Biotech/ChemLib

From: Marvich, Maria  
Sent: Thursday, January 12, 2006 2:01 PM  
To: STIC-Biotech/ChemLib  
Subject: 09/729658

*Please search SEQ ID NO:2 including interference. (391 amino acids)*

Maria Bonovich Marvich, PhD  
United States Patent and Trademark Office  
Remsen 2B84  
AU 1633  
Mail Box 2C70  
571-272-0774

RECEIVED  
JAN 12 2006  
STIC

\*\*\*\*\*

Searcher: \_\_\_\_\_  
Searcher Phone: \_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date completed: 1-18-06  
Searcher Prep Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search  
NA# \_\_\_\_\_ AA# \_\_\_\_\_  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure #: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (Specify): \_\_\_\_\_

This Page Blank (uspto)

November 2005

Published\_Applications Nucleic Acid and Published\_Applications Amino Acid database searches now generate two sets of results each. The Published\_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published\_Applications\_New databases: older published applications make up the Published\_Applications\_Main databases.

Searches run against Nucleic Acid Published\_Applications produce two sets of results, with the extensions **.rnpbm** (Published\_Applications\_NA\_Main) and **.rnpbn** (Published\_Applications\_NA\_New). Searches run against Amino Acid Published\_Applications produce two sets of results, with the extensions **.rapbm** (Published\_Applications\_AA\_Main) and **.rapbn** (Published\_Applications\_AA\_New).

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2006, 16:06:11 ; Search time 79 Seconds  
(without alignments)  
2174.647 Million cell updates/sec

Title: US-09-729-658C-2  
Perfect score: 2098  
Sequence: 1 MGFEVERRELLPAAPEER.....NMSKHTTFFGAIRLGEAPAS 391

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq 21:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*
- 9: Geneseqp2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2098	100.0	391	ABP53104	Abp53104 Human ect
2	2098	100.0	391	ABP53104	Abp53104 Human ect
3	2098	100.0	391	ABR42324	AbR42324 Human EDA
4	2098	100.0	391	ABO01928	AbO01928 Human EDA
5	2098	100.0	391	ADC35224	AdC35224 Human TNF
6	2098	100.0	391	ABW02286	AbW02286 Human EDA
7	2098	100.0	391	ADX06863	AdX06863 Cyclin-de
8	1873	89.3	377	ABO09083	AbO09083 Mouse Tab
9	1873	89.3	377	ABO01929	AbO01929 Mouse Ta
10	779	37.1	412	ADP75164	AdP75164 Fusion co
11	759	36.2	410	ADP75166	AdP75166 Fusion co
12	700	33.4	135	A4W44764	A4W44764 Human anh
13	294	14.0	357	AAR95115	AAR95115 Interveni
14	294	14.0	357	AAR95115	AAR95115 Interveni
15	294	14.0	357	AAB64008	AAB64008 CLP-CB pr
16	294	14.0	357	AAB72738	AAB72738 Repetitiv
17	286	13.6	742	ADL91066	AdL91066 Mouse col
18	282	13.4	742	AAG63347	AAG63347 Amino aci
19	282	13.4	742	ADL91057	AdL91057 Mouse col
20	282	13.4	742	ADL91062	AdL91062 Mouse col
21	280.5	13.4	1838	AAR53257	AAR53257 Human col
22	280.5	13.4	1838	ADE55566	AdE55566 Human Pro
23	280.5	13.4	1838	ADP55570	AdP55570 Human Pro
24	280.5	13.4	1838	ADP65220	AdP65220 Human alp

25	280.5	13.4	1838	9	ADV70238	Adv70238 Tumor-aes
26	280.5	13.4	1838	9	ADZ09822	AdZ09822 Human bre
27	280.5	13.4	1838	9	ADZ70492	AdZ70492 Human pro
28	277	13.2	330	2	AAW57645	Aaw57645 Collagen
29	277	13.2	408	2	AAW07539	Aaw07539 Collagen
30	277	13.2	408	4	AAAB64007	Aab64007 CLP prote
31	277	13.2	408	4	AAAB72737	Aab72737 Repetitiv
32	276	13.2	622	8	ADL91065	AdL91065 Human col
33	276	13.2	622	8	ADP56200	AdP56200 Human PRO
34	275	13.1	1466	4	AAE02534	Aae02534 Bovine al
35	275	13.1	1466	4	AAE02533	Aae02533 Bovine al
36	273.5	13.0	1014	9	ADW99577	Adw99577 Human gel
37	273	13.0	342	4	AAAM39686	Aam39686 Human pol
38	273	13.0	547	3	AAV77985	Aav77985 Human col
39	273	13.0	742	4	AAAG63346	Aag63346 Amino aci
40	273	13.0	742	5	ABBO8642	AbB08642 Human eca
41	273	13.0	742	8	ADL91071	AdL91071 Human col
42	273	13.0	742	8	ADL91064	AdL91064 Human col
43	273	13.0	1497	6	ADAB3846	AdaB3846 Human COL
44	272	13.0	1497	5	ABP43711	Abp43711 Bullous p
45	268.5	12.8	342	5	ABB84830	Abb84830 Human PRO

## ALIGNMENTS

RESULT 1  
ABP53104  
ID ABP53104 standard; protein; 391 AA.  
XX  
AC ABP53104;  
XX  
DT 11-NOV-2002 (first entry)  
XX  
DE Human ectodermal dysplasia protein (EDA) SEQ ID NO:6.  
XX

Human; adipocyte complement related protein; zacr3x2; antiinflammatory;  
KW vasotropic; antimicrobial; vulnery; antiarteriosclerotic; gene therapy;  
KW adipogenesis; gluconeogenesis; glycogenolysis; lipogenesis; ischaemia;  
KW glucose uptake; protein synthesis; thermogenesis; oxygen utilisation;  
KW inflammation; neurotransmitter; antimicrobial; infection.

OS Homo sapiens.

XX WO200246417-A2.

XX 13-JUN-2002.

XX 07-DEC-2001; 2001WO-US047348.

XX 07-DEC-2000; 2000US-0254019P.

XX (ZYMO ) ZYMOGENETICS INC.

XX Haldeman BA, Thayer EC, Sheppard PO;

XX WPI; 2002-657398/70.

PT Novel human adipocyte complement related protein, zacr3x2, useful for  
modulating energy balance in mammals, and preventing injury due to  
ischemia/inflammation.

XX Disclosure; Page 111-112; 118pp; English.

PS The present invention describes a human adipocyte complement related  
protein (I) designated zacr3x2. (I) has antiinflammatory, vasotropic,  
antimicrobial, vulnery and antiarteriosclerotic activities, and can be  
used in somatic cell gene therapy, and as a modulator of  
neurotransmission. (I) is useful for modulating energy balance in mammals  
or for protecting endothelial cells from injury, for modulating cellular  
metabolic reactions e.g. adipogenesis, gluconeogenesis, glycogenolysis,  
lipogenesis, glucose uptake, protein synthesis, thermogenesis and oxygen  
utilisation, for surgical pretreatment to prevent injury due to ischaemia

CC and/or inflammation, as neurotransmitters or as modulators of  
 CC neurotransmission e.g., modulates calcium ion concentration, muscle  
 CC contraction, hormone separation, DNA synthesis or cell growth, inositol  
 CC phosphate turnover, arachidonate release, phospholipase-C activation,  
 CC gastric emptying, human neutrophil activation or superoxide anion  
 CC production, for modulating nutrient uptake e.g. 2-deoxy-glucose uptake in  
 CC the brain, as anti-microbial agent, as cell culture reagents in vitro  
 CC studies of exogenous microorganism infection such as bacterial, viral or  
 CC fungal infection, as in vivo animal models of infection, for promoting  
 CC blood flow within the vasculature of a mammal by reducing the number of  
 CC platelets that adhere and are activated and the size of platelet  
 CC aggregates. Human zacrip2 is located on chromosome 5, more specifically  
 CC to chromosome 5p12-p13.3. The present sequence represents human  
 CC ectodermal dysplasia protein (EDA), which is given in the exemplification  
 CC of the present invention  
 XX  
 SQ Sequence 391 AA;

Query Match 100.0%; Score 2098; DB 5; Length 391;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-155;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGYPEVERRELLPAAAPRERGSGCGCGAPARAGEGNSCLLFLGFFGLSLALHLLTLCC 60  
 DB 1 MGYPEVERRELLPAAAPRERGSGCGCGAPARAGEGNSCLLFLGFFGLSLALHLLTLCC 60  
 QY 61 YLELRSELRERGAESRLGSGTPTGTSGLSLGLDPSPTITSHLGQSPKQPLEPGE 120  
 DB 61 YLELRSELRERGAESRLGSGTPTGTSGLSLGLDPSPTITSHLGQSPKQPLEPGE 120  
 QY 121 AALHSDSDQGHQALLNFFPDEKPYSEESRRVRNRKRSKNEGADGPKVKNKKKKAG 180  
 DB 121 AALHSDSDQGHQALLNFFPDEKPYSEESRRVRNRKRSKNEGADGPKVKNKKKKAG 180  
 QY 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240  
 DB 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240  
 QY 241 AGTRENQPAVHLQGGSAIQVKNDSLGGVLDNSRITMNPVKFKLHPRSSELEVLVDGT 300  
 DB 241 AGTRENQPAVHLQGGSAIQVKNDSLGGVLDNSRITMNPVKFKLHPRSSELEVLVDGT 300  
 QY 301 YFIYSQVEVYINFTDFASYEVVWDEKPFLOCTRSIETGKNTYNTCYTAGVCLLKAROKI 360  
 DB 301 YFIYSQVEVYINFTDFASYEVVWDEKPFLOCTRSIETGKNTYNTCYTAGVCLLKAROKI 360  
 QY 361 AVKVMHADISINMSKHTTFFGAIRLGEAPAS 391  
 DB 361 AVKVMHADISINMSKHTTFFGAIRLGEAPAS 391

RESULT 2  
 ABB09082  
 ID ABB09082 standard; protein; 391 AA.  
 XX  
 AC ABB09082;  
 XX  
 DT 26-JUN-2002 (first entry)  
 XX  
 DE Human EDAL-II protein SEQ ID NO:2.  
 XX  
 KW Human; mouse; EDAL-II; X-linked hypohidrotic ectodermal dysplasia; XLHED;  
 KW DL receptor; DL gene; downless gene; ectodermal dysplasia;  
 KW hair follicle growth; tooth growth; endocrine sweat gland development;  
 KW mammary epithelial tissue growth; breast epithelium.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6355782-B1.  
 XX  
 PD 12-MAR-2002.  
 XX  
 XX 29-JUN-1999; 99US-00342681.

XX 09-JUL-1998; 98US-0092279P.  
 PR 15-DEC-1998; 98US-0112366P.  
 XX  
 PA (BAYU) BAYLOR COLLEGE MEDICINE.  
 PA (UYOR-) UNIV OREGON HEALTH SCI.  
 XX  
 XX Zonana J, Ferguson BM, Headon D, Overbeek P;  
 FI WPI; 2002-360478/39.  
 XX N-PSDB; ABL51009.  
 DR  
 XX Novel purified hypohidrotic ectodermal dysplasias protein, useful for  
 PT promoting the development of hair follicles and tooth growth.  
 XX  
 PS Claim 1; Col 65-68; 86pp; English.

XX The present invention describes a purified hypohidrotic ectodermal  
 CC dysplasias protein (I) that promotes the development of hair follicles.  
 CC (I) is a promoter of hair follicle growth and tooth growth. (I) is useful  
 CC for the development of hair follicles. (I) is also useful for commercial  
 CC and clinical diagnostic testing for ectodermal dysplasia (ED). (I) is  
 CC also useful as a stimulant for hair and tooth growth, for stimulating  
 CC endocrine sweat gland development in individuals from whom the normal  
 CC sweating mechanism is compromised by disease or surgery, for stimulating  
 CC the growth of mammary epithelial tissue, either for reconstructing or  
 CC cosmetic purposes, and for promoting or maintaining differentiation of  
 CC breast epithelium. (I) is also useful for functional analysis, antibody  
 CC production and patient therapy. ABL51009 to ABL51122 and ABB09082 to  
 CC ABB09090 represent sequences used in the exemplification of the present  
 CC invention  
 XX

SQ Sequence 391 AA;

Query Match 100.0%; Score 2098; DB 5; Length 391;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-155;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGYPEVERRELLPAAAPRERGSGCGCGAPARAGEGNSCLLFLGFFGLSLALHLLTLCC 60  
 DB 1 MGYPEVERRELLPAAAPRERGSGCGCGAPARAGEGNSCLLFLGFFGLSLALHLLTLCC 60  
 QY 61 YLELRSELRERGAESRLGSGTPTGTSGLSLGLDPSPTITSHLGQSPKQPLEPGE 120  
 DB 61 YLELRSELRERGAESRLGSGTPTGTSGLSLGLDPSPTITSHLGQSPKQPLEPGE 120  
 QY 121 AALHSDSDQGHQALLNFFPDEKPYSEESRRVRNRKRSKNEGADGPKVKNKKKKAG 180  
 DB 121 AALHSDSDQGHQALLNFFPDEKPYSEESRRVRNRKRSKNEGADGPKVKNKKKKAG 180  
 QY 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240  
 DB 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240  
 QY 241 AGTRENQPAVHLQGGSAIQVKNDSLGGVLDNSRITMNPVKFKLHPRSSELEVLVDGT 300  
 DB 241 AGTRENQPAVHLQGGSAIQVKNDSLGGVLDNSRITMNPVKFKLHPRSSELEVLVDGT 300  
 QY 301 YFIYSQVEVYINFTDFASYEVVWDEKPFLOCTRSIETGKNTYNTCYTAGVCLLKAROKI 360  
 DB 301 YFIYSQVEVYINFTDFASYEVVWDEKPFLOCTRSIETGKNTYNTCYTAGVCLLKAROKI 360  
 QY 361 AVKVMHADISINMSKHTTFFGAIRLGEAPAS 391  
 DB 361 AVKVMHADISINMSKHTTFFGAIRLGEAPAS 391

RESULT 3  
 ABR42324  
 ID ABR42324 standard; protein; 391 AA.  
 XX  
 AC ABR42324;  
 XX







PR 07-JUL-2000; 2000US-0216879P.  
PR 26-MAR-2001; 2001US-0278449P.  
PR 06-JUL-2001; 2001US-00899059.  
PR 24-AUG-2001; 2001US-0314381P.  
PR 27-DEC-2001; 2001US-0336695P.  
PR 23-AUG-2002; 2002US-0022629A.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX  
PI Yu G, Ni J, Rosen CA, Zhang J, Wei P;  
XX  
XX WPI; 2003-852773/79.  
DR N-PSDB; AAD63922.  
XX  
PT Use of tumor necrosis factor gamma-beta antagonists for treating or  
PT ameliorating a disease or disorders of the gastrointestinal tract, e.g.  
PT inflammatory bowel disease, Crohn's disease or ulcerative colitis.  
XX  
XX Disclosure; Page 156-157; Opp; English.  
XX  
XX The invention relates to methods and compositions for treating or  
CC ameliorating a disease or disorder of the gastrointestinal tract. The  
CC method involves administering a composition comprising tumour necrosis  
CC factor (TNF)-gamma-beta and its receptors DR3 and TR6 to a person with,  
CC or suspected of having the disease or disorder. The antagonist of TNF-  
CC gamma-beta is useful for treating or ameliorating a gastrointestinal  
CC tract disease or disorder, specifically an inflammatory bowel disease,  
CC e.g. Crohn's disease or ulcerative colitis. The present sequence is human  
CC EDA protein. This sequence is used to illustrate the method of the  
XX  
XX Sequence 391 AA;  
Query Match 100.0%; Score 2098; DB 7; Length 391;  
Best Local Similarity 100.0%; Pred. No. 5.9e-155;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MGYPEVERRELLPAAAPRERGSGCGCGGAPARAGEGNSCLLFLGFFGLSLALHLTLCC 60  
Db 1 MGYPEVERRELLPAAAPRERGSGCGCGGAPARAGEGNSCLLFLGFFGLSLALHLTLCC 60  
QY 61 YLELSELRRERGAESRLGGSGTPTGTSGLSSGLGLDPSPTSHLGQSPKQOPLPG 120  
Db 61 YLELSELRRERGAESRLGGSGTPTGTSGLSSGLGLDPSPTSHLGQSPKQOPLPG 120  
QY 121 AALHSDSQDHQWALLNFFPPDEKPYSEESRVRNRKRSKNEGADGVPVKNKKKAG 180  
Db 121 AALHSDSQDHQWALLNFFPPDEKPYSEESRVRNRKRSKNEGADGVPVKNKKKAG 180  
QY 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240  
Db 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240  
QY 241 AGTRENQPAVHLQGGSAIQVKNDSLGGVLDNSRITWNPVKVFLHPRSGELEVLDGT 300  
Db 241 AGTRENQPAVHLQGGSAIQVKNDSLGGVLDNSRITWNPVKVFLHPRSGELEVLDGT 300  
QY 301 YFTYSQVEVYVYINFTDFASVEVVVDEKPFLOCTRSIETGKNTNTCTTAGVCLLKARQKI 360  
Db 301 YFTYSQVEVYVYINFTDFASVEVVVDEKPFLOCTRSIETGKNTNTCTTAGVCLLKARQKI 360  
QY 361 AVKQVHADISINMSKHTTFFGAILRGEAPAS 391  
Db 361 AVKQVHADISINMSKHTTFFGAILRGEAPAS 391  
RESULT 7  
ADX06863  
ID ADX06863 standard; protein; 391 AA.  
XX  
AC ADX06863;  
XX  
DT 21-APR-2005 (first entry)

XX Cyclin-dependent kinase modulation biomarker SEQ ID NO 1428.  
DE cytostatic; cyclin-dependent kinase; cdk; biomarker.  
XX  
KW Homo sapiens.  
OS  
XX WO2005012875-A2.  
PN 10-FEB-2005.  
PD 29-JUL-2004; 2004WO-US024424.  
PF 29-JUL-2003; 2003US-0490890P.  
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
PA  
XX Li M, Rupnow BA, Webster KR, Jackson DG, Wong TW;  
PI WPI; 2005-163068/17.  
XX N-PSDB; ADX06862.  
DR Biomarkers useful for predicting or determining the response of a mammal  
XX to a cancer treatment comprising administration of a modulator of cyclin-  
XX dependent kinase activity.  
XX Claim 5; SEQ ID NO 1428; 141pp; English.  
XX  
XX This invention describes a novel method of predicting or determining  
CC whether a mammal will respond or is responding to an anti-cancer agent  
CC that modulates cyclin-dependent kinase (cdk) activity. The method  
CC comprises measuring the level of one or more biomarkers selected from  
CC 2774 biomarkers given in the specification (nucleotide sequence SEQ ID  
CC NO:1246 (Genbank EST W28729) is especially preferred). The method of the  
CC invention is utilized in a kit for determining or predicting whether  
CC patient would be susceptible or resistant to treatment by an agent  
CC modulating cdk activity. The invention also describes a method for  
CC utilizing individualized genetic profiles for treating diseases and  
CC disorders based on patient's response and molecular level, specialized  
CC microarrays comprising the biomarkers described, antibodies directed  
CC against the biomarkers and a cell culture model to identify biomarkers.  
CC The cdk modulator is preferably N-5-[(5-(1,1-Dimethylethyl)-2-oxazolyl)methyl]thio]-2-thiazolyl-4-piperidine carboxamide, 0.5-L-  
CC tartaric acid salt. Note: The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences. This  
CC sequence represents a biomarker used in the method of the invention.  
XX  
XX Sequence 391 AA;  
Query Match 100.0%; Score 2098; DB 9; Length 391;  
Best Local Similarity 100.0%; Pred. No. 5.9e-155;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MGYPEVERRELLPAAAPRERGSGCGCGGAPARAGEGNSCLLFLGFFGLSLALHLTLCC 60  
Db 1 MGYPEVERRELLPAAAPRERGSGCGCGGAPARAGEGNSCLLFLGFFGLSLALHLTLCC 60  
QY 61 YLELSELRRERGAESRLGGSGTPTGTSGLSSGLGLDPSPTSHLGQSPKQOPLPG 120  
Db 61 YLELSELRRERGAESRLGGSGTPTGTSGLSSGLGLDPSPTSHLGQSPKQOPLPG 120  
QY 121 AALHSDSQDHQWALLNFFPPDEKPYSEESRVRNRKRSKNEGADGVPVKNKKKAG 180  
Db 121 AALHSDSQDHQWALLNFFPPDEKPYSEESRVRNRKRSKNEGADGVPVKNKKKAG 180  
QY 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240  
Db 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240  
QY 241 AGTRENQPAVHLQGGSAIQVKNDSLGGVLDNSRITWNPVKVFLHPRSGELEVLDGT 300  
Db 241 AGTRENQPAVHLQGGSAIQVKNDSLGGVLDNSRITWNPVKVFLHPRSGELEVLDGT 300







Db 121 AALHSDSDQGHQ 132

RESULT 13

AAAR95115

ID AAR95115 standard; protein; 357 AA.

XX AC AAR95115;

XX 25-MAR-2003 (revised)

XX 03-FEB-1997 (first entry)

XX

XX Intervening sequence contg. collagen like protein (CLP)-CB.

DE DE

XX Polymer; repeat unit; natural collagen; intervening oligopeptide; fibre; film; membrane; emulsion; coating; collagen like protein; specific binding material; catalyst; purification agent; composite; laminated; adhesive; cell growth surface; affinity column; biological material support; wound dressing; in vivo prothesis.

XX OS Synthetic.

XX US5514581-A.

XX 07-MAY-1996.

XX 06-NOV-1990; 90US-00609716.

XX 04-NOV-1986; 86US-00927258.

XX 29-OCT-1987; 87US-00114618.

XX 09-NOV-1988; 88US-00269429.

XX 07-NOV-1989; 89WO-00505016.

(PROT-) PROTEIN POLYMER TECHNOLOGIES INC.

XX Cappelletto J, Ferrari FA;

XX WPI; 1996-238772/24.

XX DNA encoding protein contg. repeated fibroin derived segments - linked by oligopeptide with cell adhesion properties useful, e.g. in wound dressings.

XX Example 3; Col 101-104; 71pp; English.

XX Novel DNA sequence encodes a polymer comprising segments of repeating units of 3-9 amino acids from natural collagen, i.e. the present intervening sequence contg. collagen like protein (CLP)-CB, able to assemble into aligned structures formable into articles. The polymer comprises at least 2 segments joined by an unaligned intervening oligopeptide, other than the repeating unit. The polymer can be used to make fibres, films, membranes, emulsions, coatings, etc., useful as, e.g. specific binding materials, catalysts, purific. agents, composites, laminates, adhesives, cell growth surfaces, affinity columns and supports for biological materials. Typical applications include wound dressings, and in vivo protheses. The polymer produces articles with good mechanical properties, and the intervening oligopeptide can provide a ligand for binding a mol., antibody, etc., or a chemically reactive site for coupling to proteins, etc. . (Updated on 25-MAR-2003 to correct PF field.)

XX SQ Sequence 357 AA;

Query Match 14.0%; Score 294; DB 2; Length 357;

Best Local Similarity 32.4%; Pred. No. 2e-14;

Matches 69; Conservative 9; Mismatches 71; Indels 64; Gaps 3;

QY 71 ERGAERLGGSGTPTGTSGLSLGLGLDPPSP--ITSHLGQSPKQOPLPEPEALHSDSQ 128

Db 71 DRDAGPKGADSGFGAGPVGPGPPGPGPGAGCPGPGPPGPGPLGPKG 130

QY 129 DGHQALLNPPFFDEKPYSEESRRVRRNKRKSKSNEGADGPVKNKKKKKAGPAGPGNGPP 188

Db 131 D-----RGDAGPKGADSGFGAGPVGPGPPGPGPGAGCPGPGPPGPGPP 165



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2006, 16:08:46 ; Search time 16 Seconds  
(without alignments)  
2351.297 Million cell updates/sec

Title: US-09-729-658C-2

Perfect score: 2098

Sequence: 1 MGYPEVERRELLPAAAPRER.....NMSKHTTFFGAILGEAPAS 391

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 80:\*

1: Pir1:\*

2: Pir2:\*

3: Pir3:\*

4: Pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	280.5	13.4	1038	1 CGHU1V	collagen alpha 1(V)
2	277	13.2	1464	2 S59856	collagen alpha 1(I)
3	275.5	13.1	1433	2 A46053	bullous pemphigoid
4	275	13.1	1049	1 CGB07S	collagen alpha 1(I)
5	273.5	13.0	1843	2 S18803	collagen alpha 1(V)
6	273	13.0	1532	2 A61262	collagen alpha 1(X)
7	268.5	12.8	742	2 JCT595	scavenger receptor
8	267.5	12.8	316	2 T20497	hypothetical prote
9	267	12.7	886	2 I50694	collagen alpha 1(I)
10	267	12.7	1315	2 A56101	collagen alpha 1(X)
11	267	12.7	1774	2 B56101	collagen alpha 1(X)
12	266	12.7	779	1 CGB01S	collagen alpha 1(I)
13	265.5	12.7	1464	1 CGHU1S	collagen alpha 1(I)
14	265	12.6	1763	2 S16366	collagen alpha 2(I)
15	264	12.6	1690	1 CGHU1B	collagen alpha 4(I)
16	263.5	12.6	1146	2 A38587	collagen, cornea-s
17	263.5	12.6	1419	2 A41182	collagen alpha 1(I)
18	263.5	12.6	1487	2 B41182	collagen alpha 1(I)
19	263	12.5	1366	1 CGHU2S	collagen alpha 2(I)
20	260.5	12.4	675	2 S20819	collagen alpha 3(I)
21	259.5	12.4	302	2 T15936	hypothetical prote
22	259.5	12.4	674	2 S23297	collagen alpha 1(X)
23	258.5	12.3	1373	1 A43291	collagen alpha 2(I)
24	258	12.3	674	2 S13501	collagen alpha 1(X)
25	258	12.3	680	1 CGHU1D	collagen alpha 1(X)
26	258	12.3	1142	2 JX0369	collagen alpha 1(X)
27	257.5	12.3	312	2 T25048	hypothetical prote
28	257.5	12.3	1453	2 S21626	collagen alpha 1(I)
29	256.5	12.2	688	2 A53330	collagen alpha 2(I)

ALIGNMENTS

RESULT 1

CGHU1V  
collagen alpha 1(V) chain precursor - human  
N:Alternate names: procollagen alpha 1(V) chain  
C:Species: Homo sapiens (man)  
C>Date: 22-Nov-1993 #sequence revision 03-Oct-1995 #text change 09-Jul-2004  
C:Accession: S18802; S16024; A61142; S11303; S03978; S43642; S58665  
J:Greenpan, D.S.; Cheng, W.; Hoffman, G.G.  
J: Biol. Chem. 266, 24727-24733, 1991  
A:Title: The pro-alpha1(V) collagen chain. Complete primary structure, distribution of  
A:Reference number: S18802; MUID:92105142; PMID:1722213  
A:Accession: S18802  
A:Molecule type: mRNA  
A:Residues: 1-1838 <GR>  
A:Cross-references: UNIPROT:Q15094; UNIPARC:UPI000006E982; GB:M76729; NID:G189519; PIDN  
R:Takahara, K.; Sato, Y.; Okazawa, K.; Okamoto, N.; Noda, A.; Yaoi, Y.; Kato, I.  
J. Biol. Chem. 266, 13124-13129, 1991  
A:Title: Complete primary structure of human collagen alpha-1(V) chain.  
A:Reference number: S16024; MUID:91302336; PMID:2071595  
A:Accession: S16024  
A:Molecule type: protein  
A:Residues: 1-81, 'QL', '84-389, 'A', '391-676, 'K', '678-1294, 'PS', '1297, 'RS', '1300-1553, 'R', '1555  
A:Cross-references: UNIPARC:UPI0000126D1D; GB:D90279; NID:G219509; PIDN:BA414323.1; PID  
A:Note: Parts of this sequence were determined by protein sequencing  
R:Yaoi, Y.; Hashimoto, K.; Takahara, K.; Kato, I.  
Exp. Cell Res. 194, 180-185, 1991  
A:Title: Insulin binds to type V collagen with retention of mitogenic activity.  
A:Reference number: A61142; MUID:91224163; PMID:1709100  
A:Accession: A61142  
A:Molecule type: protein  
A:Residues: 823-824, 'X', '826-842 <YAO>  
A:Cross-references: UNIPARC:UPI0000173BB7  
A:Note: The residue designated 'X' is probably glycosylated hydroxylysine; this cyanoge  
R:Yaoi, Y.; Hashimoto, K.; Koitabashi, H.; Takahara, K.; Ito, M.; Kato, I.  
Biochim. Biophys. Acta 1035, 139-145, 1990  
A:Title: Primary structure of the heparin-binding site of type V collagen.  
A:Reference number: S11303; MUID:90366601; PMID:2203476  
A:Accession: S11303  
A:Molecule type: protein  
A:Residues: 823-824, 'X', '826-848, 'I', '850-851, 'P', '853, 'PR', '856-893, 'D', '895-932, 'X', '934-935  
A:Cross-references: UNIPARC:UPI0000173BB8  
A:Note: The residues designated 'X' are probably glycosylated hydroxylysine; this sequ  
R:Seyer, J.M.; Kang, A.H.  
Arch. Biochem. Biophys. 271, 120-129, 1989  
A:Title: Covalent structure of collagen: amino acid sequence of three cyanogen bromide-  
A:Reference number: S03978; MUID:89227189; PMID:2496661  
A:Accession: S03978  
A:Molecule type: protein  
A:Residues: 621-640, 'G', '642-649, 'L', '651-662, 'E', '664-667, 'Q', '669-676, 'Q', '678-683, 'P', '685  
A:Cross-references: UNIPARC:UPI0000173BB9  
A:Note: there are a number of inconsistencies between the sequences in figures 6 and 7;  
<SB>

collagen alpha 1(I)  
collagen alpha 1(I)  
collagen alpha 3(I)  
collagen alpha 2(I)  
collagen alpha 2(I)  
collagen alpha 1(V)  
collagen alpha 1(X)  
collagen alpha 1(I)  
collagen alpha 1(I)  
collagen alpha 1(I)  
hypothetical prote  
hypothetical prote  
ORF2 protein - Sal  
collagen alpha 3(I)  
collagen alpha 1(X)





















A;Reference number: A35336; MUID:90252792; PMID:2339700  
A;Accession: A35336  
A;Molecule type: mRNA  
A;Residues: 710-720,'E',722-737,'E',739-745 <WAL>  
A;Cross-references: UNIPARC:UPI0000173B4B  
A;Note: the authors translated the codons CAG for 721 and CGT for 738 as Glu  
R;Forlino, A.; Zolezzi, P.; Valli, M.; Pignatti, P.C.; Motte  
Hum. Mol. Genet. 3, 2201-2206, 1994  
A;Title: Severe (type III) osteogenesis imperfecta due to glycine substitutions in the  
A;Reference number: I54365; MUID:95187161; PMID:7881420  
A;Accession: I54365  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 746-766,'S',768-781 <FOR>  
A;Cross-references: UNIPARC:UPI0000173B4B; GB:L47667; NID:G1009093; PIDN:AAB59576.1; PID  
R;Chesler, S.D.; Wallis, G.A.; Byers, P.H.  
J. Biol. Chem. 268, 18218-18225, 1993  
A;Title: Mutations in the carboxyl-terminal propeptide of the pro alpha 1(I) chain of ty  
A;Reference number: A47426; MUID:93352646; PMID:8349697  
A;Accession: A47426  
A;Molecule type: mRNA  
A;Residues: 1179-1276,'H',1278-1336,1339-1387,'R',1389-1464 <CHE>  
A;Cross-references: UNIPARC:UPI0000173B4B; GB:S64596; NID:G407589; PIDN:AAB27856.1; PID:  
A;Note: sequence extracted from NCBI backbone (NCBIN:136444, NCBI:P136445)  
A;Note: does not represent an experimentally determined sequence but three different mut  
A;Accession: B47426  
A;Molecule type: mRNA  
A;Residues: 1179-1464 <CH4>  
A;Cross-references: UNIPARC:UPI0000173B4C  
A;Experimental source: normal dermal fibroblast culture  
A;Accession: C47426  
A;Molecule type: mRNA  
A;Residues: 1179-1276,'H',1278-1464 <CH5>  
A;Cross-references: UNIPARC:UPI0000173B4D  
A;Experimental source: fetal cell 86-237  
A;Accession: D47426  
A;Molecule type: mRNA  
A;Residues: 1179-1336,1339-1464 <CH6>  
A;Cross-references: UNIPARC:UPI0000173B4E  
A;Experimental source: fetal cell 86-146  
A;Accession: E47426  
A;Molecule type: mRNA  
A;Residues: 1179-1387,'R',1389-1464 <CH7>  
A;Cross-references: UNIPARC:UPI0000173B4F  
A;Experimental source: fetal cell 88-251  
R;Cohn, D.H.; Apone, S.; Eyre, D.R.; Starman, B.J.; Andraessen, P.; Charbonneau, H.; Nic  
J. Biol. Chem. 263, 14605-14607, 1988

Query Match 12.7%; Score 265.5; DB 1; Length 1464;  
Best Local Similarity 29.7%; Pred. No. 1.3e-09;  
Matches 78; Conservative 16; Mismatches 80; Indels 89; Gaps 9;  
QY 13 PAAAPRRGSGQ-----CGCGGAPARAGE-----GNSCLLPLGFGGLSLALHLTLCCY 61  
DB 774 PAGAPDGGEGSPGAGTGAAGPGRGPPGPA-----GFGA----- 815  
QY 62 LELSELRRERGAESRLGSGTPTGTSGLSSGLGDDPSITSHLGQSPKQOPLPEGEA 121  
DB 816 -----PPGADGQPGAKGEGPDAGAKDAGPPGAPG-----AGPGPIGNVGPAG- 861  
QY 122 ALHSDSQDHQWALLNFFPDEKPYSEBSRRVRNRKSKNEGADGVPVKNKKKGKAGP 181  
DB 862 -----KGARSAGPPGATGPFGAAGRVGP 885  
QY 182 PGPNPPGPPGPPG-----GGPPGPIGPIGTVMGPPGPPGPPGPPGPPGPPGPPG 235  
DB 886 PGPSGNAGPPGPPGAGKEGKGKPGTGPAGRPGE-----VGPPGPPGAGKGSFGADGPA 943  
QY 236 GAADKAGTR--ENQPAVVHLOQQ 256  
DB 944 GAGTGTGGIAGQAGVWGLPQ 966

## RESULT 14

S16366  
Collagen alpha 2(IV) chain precursor - pig roundworm  
C;Species: Ascaris suum (pig roundworm)  
C;Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 09-Jul-2004  
C;Accession: S16366  
R;Pettitt, J.; Kingston, I.B.  
J. Biol. Chem. 266, 16149-16156, 1991  
A;Title: The complete primary structure of a nematode alpha-2(IV) collagen and the parti  
A;Reference number: S16366; MUID:91340768; PMID:1714907  
A;Accession: S16366  
A;Molecule type: mRNA  
A;Residues: 1-1763 <JBI>  
A;Cross-references: UNIPROT:P27393; UNIPARC:UPI0000126D40; GB:M67507; NID:G159648; PIDN:  
C;Genetics: 229/3; 266/3; 305/3; 360/3; 424/1; 489/1; 548/1; 656/3; 790/1; 891/1; 963/1;  
A;Introns: 229/3; 266/3; 305/3; 360/3; 424/1; 489/1; 548/1; 656/3; 790/1; 891/1; 963/1;  
C;Superfamily: collagen alpha 1(IV) chain  
C;Keywords: alternative splicing; basement membrane; cell binding; coiled coil; disulfid  
F;1-26/Domain: signal sequence #status predicted <SIG>  
F;27-1763/Product: collagen alpha 2(IV) chain #status predicted <MAT>  
F;27-42/Domain: non-collagenous NHI #status predicted <NHI>  
F;43-1529/Domain: collagenous #status predicted <COL>  
F;197-199/Region: cell attachment (R-G-D) motif  
F;1530-1763/Domain: carboxyl-terminal nonhelical, NCI #status predicted <NC1>  
F;1530-1638/Domain: repeat NCI #status predicted <NC11>  
F;1639-1763/Domain: repeat NCI #status predicted <NC12>  
F;31,34,39,41,536,539/Disulfide bonds: interchain #status predicted  
F;126/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;1593-1599,1702-1709/Disulfide bonds: #status predicted

Query Match 12.6%; Score 265; DB 2; Length 1763;  
Best Local Similarity 34.3%; Pred. No. 1.6e-09;  
Matches 85; Conservative 15; Mismatches 82; Indels 66; Gaps 10;

QY 13 PAAAPRRGSGQ-----CGCGGAPARAGEGNSCLLPLGFGGLSLALHLTLCCYLELR 65  
DB 256 PEGAGPMKGEKGDPEAGPRGFTGTVAGQP-----GLPGMKGKGL----- 298  
QY 66 SELRRERGAESRLGSGTPTGTSGLSSGLGDDPSITSHLGQSPKQOPLPEGEAALHS 125  
DB 299 SGAPAGPRGKGRPLGPPGPKGD-----RGLD-----GLPGVGLPQKGEAGF-- 343  
QY 126 DSQDHQWALLNFFPDEKPYSEBSRRVRNRKSKNEGADGVPVKNKKKGKAGPPGN 185  
DB 344 PGRDCAKAR-----GPPGPPGGEFS-----DGPPGPPGLPGREGQPGPPGAD 387  
QY 186 GPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 236  
DB 388 GYPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 445  
QY 237 AADKAGTR 244  
DB 446 AKGPPGPR 453

## RESULT 15

CGHULB  
collagen alpha 4(IV) chain precursor - human  
N;Alternate names: procollagen alpha 4(IV) chain  
C;Species: Homo sapiens (man)  
C;Date: 06-Feb-1995 #sequence\_revision 03-Oct-1995 #text\_change 09-Jul-2004  
C;Accession: A55360; S36854; S28777  
R;Leinonen, A.; Mariyama, M.; Mochizuki, T.; Tryggvason, K.; Reiders, S.T.  
J. Biol. Chem. 269, 26172-26177, 1994  
A;Title: Complete primary structure of the human type IV collagen alpha4(IV) chain. Com  
A;Reference number: A55360; MUID:95014445; PMID:7523402  
A;Accession: A55360  
A;Status: nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-1690 <LEI>  
A;Cross-references: UNIPROT:P53420; UNIPARC:UPI0000000769; GB:X81053; NID:G574805; PIDN:  
R;Sugimoto, M.; Ohashi, T.; Yoshioka, H.; Matsuo, N.; Ninomiya, Y.

FEBS Lett. 330, 122-128, 1993  
A;Title: cDNA isolation and partial gene structure of the human alpha-4 (IV) collagen chain  
A;Reference number: S36854; MUID:93374047; PMID:8365481  
A;Accession: S36854  
A;Molecule type: DNA; mRNA  
A;Residues: 1219-1658, 'FE', 1661-1690 <SUG>  
A;Cross-references: UNIPARC:UPI000016A432; DDBJ:DJ17391; NID:g440365; PIDN:BA04214.1; PI  
R;Kamagata, Y.; Mattei, M.G.; Nimomiya, Y.  
J. Biol. Chem. 267, 23753-23758, 1992  
A;Title: Isolation and sequencing of cDNAs and genomic DNAs encoding the alpha4 chain of  
A;Reference number: S28777; MUID:93054733; PMID:1429714  
A;Accession: S28777  
A;Molecule type: DNA  
A;Residues: 1407-1424, 'G', 1426-1430, 'A', 1432-1439, 'L', 1441-1507 <KAM>  
A;Cross-references: UNIPARC:UPI0000173BFC; GB:L01475; GB:L01476  
A;Note: the codons given for 1438-Asp (GAG) and 1443-Gly (GCA) are inconsistent with the  
C;Comment: Prolines and lysines at the third position of the tripeptide repeating unit  
ed and subsequently O-glycosylated.  
C;Genetics:  
A;Gene: GDB:COL4A4  
A;Cross-references: GDB:132673; OMIM:120131  
A;Map position: 2q35-2q37  
A;Introns: 39/1; 1406/1; 1445/1; 1508/1; 1603/3 #status incomplete  
A;Note: the alpha 3(IV) and alpha 4(IV) chain genes are encoded on opposite strands with  
C;Complex: this minor type IV collagen is thought to form a heterotrimer of two alpha 3  
among trimer amino-terminal domains (with disulfide and desmosine cross-links), dimeric a  
er associations in the interrupted helical domain (with disulfide and desmosine cross-li  
C;Function:  
A;Description: minor structural component of extracellular basement membrane in kidney  
C;Superfamily: collagen alpha 1(IV) chain  
C;Keywords: basement membrane; coiled coil; extracellular matrix; glycoprotein; hydroxyl  
F;1-38/Domain: signal sequence #status predicted <SIG>  
F;39-1690/Product: collagen alpha 4(IV) chain #status predicted <MAT>  
F;39-61/Domain: amino-terminal nonhelical, NH1 <NH1>  
F;62-1466/Region: interrupted helical  
F;94-96/Region: cell attachment (R-G-D) motif  
F;145-147/Region: cell attachment (R-G-D) motif  
F;189-191/Region: cell attachment (R-G-D) motif  
F;310-312/Region: cell attachment (R-G-D) motif  
F;724-726/Region: cell attachment (R-G-D) motif  
F;785-787/Region: cell attachment (R-G-D) motif  
F;989-991/Region: cell attachment (R-G-D) motif  
F;1212-1214/Region: cell attachment (R-G-D) motif  
F;1467-1690/Domain: carboxyl-terminal nonhelical, NC1 <NC1>  
F;1471-1569/Domain: collagen IV carboxyl-terminal repeat <CT1>  
F;1579-1686/Domain: collagen IV carboxyl-terminal repeat <CT2>  
F;47,52,55,57,266,400,460,492,494,668,790,828,1095,1131,1294,1317,1375,1407/Disulfide bc  
F;142,669/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;1480-1566,1513-1569/Disulfide bonds: (or 1480-1569, 1513-1566) #status predicted  
F;1525-1531,1634-1641/Disulfide bonds: #status predicted  
F;1588-1683,1622-1686/Disulfide bonds: (or 1588-1686, 1622-1683) #status predicted

Query Match 12.6%; Score 264; DB 1; Length 1690;  
Best Local Similarity 27.4%; Pred. No. 1.8e-09;  
Matches 96; Conservative 18; Mismatches 96; Indels 140; Gaps 15;

Qy	2	GYPEVERELLPAAPRERSGQ-CGCGGAPARAGEGNSCLFLGFGGLSLALHLLTCC 60
Db	1008	GPFGHKE-----PGKQGPFGPPGPGSTGLRG-----FIGFGLP----- 1047
Qy	61	YLELRSELRERGAESRLGGSGTPTGTLSSLGGLDPSITSHLQGPSKQPLRPG- 119
Db	1048	-----GQQGERGSPGPGFGIGDARGPKGNKGDPAHFPGPKGPGSPGC 1095
Qy	120	-----EALHSDSQD-----HOMALLNFFPPDEK----- 144
Db	1096	PGHFGASGEQLPGTQGRGFRGPPGPPGSSGPPGCDGHGMPGLR-GQPGEMGDPPRG 1154
Qy	145	-----PYSEESRRVRNRNKRKSNEGAD-----GPV-----KNKK-- 174
Db	1155	LQDPGIPGPGIKGPSGLNGLHGLKQKGTKGASGLHDVGPVGPVGLKGERGD 1214

Qy	175	-----KKKA--GPPGNGPPG-----GPPGPGPP-----GI 202
Db	1215	PGSPGISPPGPRGKGGPPGPGSSGPPGAGATGRAPKDIIPDPGPPDQGPDPGRGA 1274
Qy	203	PGIPGIPGTTVM-----GPPGPPGPPGPPGPGPGPGSGAADKAGTR 244
Db	1275	PGPGLPGSVDLLRGEFGDGLPGPFGPPGPPGPGPGPGKGFPGCDGKDGQK 1324

Search completed: January 13, 2006, 16:15:22  
Job time : 17 secs



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2006, 16:07:11 ; Search time 73 Seconds  
(without alignments)  
3778.926 Million cell updates/sec

Title: US-09-729-658C-2

Perfect score: 2098

Sequence: 1 MGYPEVERRELLPAAAPRRER.....NMSKHTTFFGAILGEPAS 391

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot\_05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2098	100.0	391	1	EDA_HUMAN	Q92838 homo sapien
2	1972	94.0	391	1	EDA_BOVIN	Q98653 bos taurus
3	1965	93.7	391	1	EDA_MOUSE	O54893 mus musculus
4	1890.5	90.1	386	2	Q58P77_CANFA	Q58P77 canis faml
5	1441.5	68.7	356	2	Q58F25_CHICK	Q58F25 gallus gall
6	955.5	45.5	367	2	Q50D54_GASAC	Q50D54 gasterosteu
7	887	42.3	178	2	Q5JS00_HUMAN	Q5JS00 homo sapien
8	809	38.6	207	2	Q4SJR1_TETNG	Q4SJR1 tetraodon n
9	286	13.6	742	2	Q8VIF6_MOUSE	Q8VIF6 mus musculus
10	284	13.5	998	2	Q8CFM4_MOUSE	Q8CFM4 mus musculus
11	282	13.4	742	2	Q8K4Q8_MOUSE	Q8K4Q8 mus musculus
12	282	13.4	742	2	Q8C979_MOUSE	Q8C979 mus musculus
13	280.5	13.4	1792	2	Q59EE7_HUMAN	Q59EE7 homo sapien
14	280.5	13.4	1838	1	Q5A1_HUMAN	P20908 homo sapien
15	280.5	13.4	1838	2	Q15094_HUMAN	Q15094 homo sapien
16	280.5	13.4	1838	2	Q5SUX4_HUMAN	Q5SUX4 homo sapien
17	280.5	13.4	1840	2	Q59IP3_PIG	Q59IP3 sus scrofa
18	280	13.3	284	2	Q25581_9BILA	Q25581 teladorsagi
19	278	13.3	742	2	Q4V885_RAT	Q4V885 rattus norv
20	277.5	13.2	1838	2	Q88207_MOUSE	Q88207 mus musculus
21	277.5	13.2	1840	2	Q9J103_RAT	Q9J103 rattus norv
22	276	13.2	622	2	Q9BY85_HUMAN	Q9BY85 homo sapien
23	276	13.2	1315	2	Q8QHL9_XENLA	Q8QHL9 xenopus lae
24	275.5	13.1	1470	1	COHAI1_MOUSE	Q07563 mus musculus
25	275	13.1	1049	1	CO3A1_BOVIN	P04258 bos taurus
26	275	13.1	1431	1	COHAI1_MESAU	Q9JMB4 mesocricetu
27	273.5	13.0	1840	2	Q60467_CRILLO	Q60467 cricetus
28	273	13.0	742	2	Q5KU26_HUMAN	Q5KU26 homo sapien
29	273	13.0	742	2	Q8WZ44_HUMAN	Q8WZ44 homo sapien
30	273	13.0	1497	1	COHAI1_HUMAN	Q9UMD9 homo sapien
31	272	13.0	284	2	Q25582_9BILA	Q25582 teladorsagi

32 272 13.0 1415 2 Q5JV36\_HUMAN Q5JV36 homo sapien  
33 272 13.0 1802 2 Q17163\_BRUMA Q17163 bugia mala  
34 271.5 12.9 1344 2 Q93419\_CHICK Q93419 gallus gall  
35 271 12.9 326 2 Q5C1A3\_SCHJA Q5C1A3 schistosoma  
36 270 12.9 1222 2 Q8K173\_MOUSE Q8K173 mus musculus  
37 270 12.9 1464 1 Q3A1\_MOUSE P08121 mus musculus  
38 270 12.9 1464 2 Q8BLW4\_MOUSE Q8BLW4 mus musculus  
39 270 12.9 1464 2 Q7TT32\_MOUSE Q7TT32 mus musculus  
40 270 12.9 1467 2 Q5DTG2\_MOUSE Q5DTG2 mus musculus  
41 268.5 12.8 742 2 Q9BYH7\_HUMAN Q9BYH7 homo sapien  
42 268.5 12.8 742 2 Q8P9P2\_HUMAN Q8P9P2 homo sapien  
43 268.5 12.8 1258 2 Q8AW11\_BRARE Q8AW11 brachydantio  
44 268.5 12.8 1835 2 Q9IAU4\_CHICK Q9IAU4 gallus gall  
45 267.5 12.8 316 2 Q19111\_CAEEL Q19111 caenorhabdi

#### ALIGNMENTS

#### RESULT 1

EDA\_HUMAN STANDARD; PRT: 391 AA  
AC Q92838; O75910; Q5JUM7; Q9UP77; Q9Y6L0; Q9Y6L1; Q9Y6L2; Q9Y6L3;  
AC Q9Y6L4;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Ectodysplasin A (Ectodermal dysplasia protein) (EDA protein)  
DE [Contains: Ectodysplasin A, membrane form; Ectodysplasin A, secreted form].  
GN Name=EDA; Synonyms=ED1;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA / MRNA] (ISOFORM I), AND VARIANTS EDA  
RP HIS-61 AND LEU-69.  
RC TISSUE=Sweat gland;  
RX MEDLINE=96331280; PubMed=8696334;  
RA Kere J., Srivastava A.K., Montonen O., Zonana J., Thomas N.S.T.,  
RA Ferguson B.M., Munoz P., Morgan D., Clarke A., Baybayan P., Chen E.Y.,  
RA Ezer S., Saarialho-Kere U., la Chapelle A., Schlesinger D.;  
RT "X-linked anhidrotic (hypohidrotic) ectodermal dysplasia is caused by  
RL mutation in a novel transmembrane protein.";  
RL Nat. Genet. 13:409-416(1996).  
RN [2]  
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM A1), AND VARIANTS EDA.  
RC TISSUE=Liver;  
RX MEDLINE=98349961; PubMed=9683615;  
RA Monreal A.W., Zonana J., Ferguson B.M.;  
RT "Identification of a new splice form of the EDAl gene permits  
RT detection of nearly all X-linked hypohidrotic ectodermal dysplasia  
RT mutations.";  
RL Am. J. Hum. Genet. 63:380-389(1998).  
RN [3]  
RP NUCLEOTIDE SEQUENCE (ISOFORMS A1; A2; B; C; D; E AND F), AND VARIANTS  
RP EDA.  
RX MEDLINE=98409495; PubMed=9736768; DOI=10.1093/hmg/7.11.1661;  
RA Bayes M., Hartung A.J., Ezer S., Piepa J., Theeleff I.,  
RA Srivastava A.K., Kere J.;  
RT "The anhidrotic ectodermal dysplasia gene (EDA) undergoes alternative  
RT splicing and encodes ectodysplasin-A with deletion mutations in  
RL collagenous repeats.";  
RL Hum. Mol. Genet. 7:1661-1669(1998).  
RN [4]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RX PubMed=15772651; DOI=10.1038/nature03440;  
RA Ross M.T., Grafham D.V., Coffey A.J., Scherer S., McLeay K., Muzny D.,  
RA Platzer M., Howell G.R., Burrows C., Bird C.P., Frankish A.,  
RA Lovell F.L., Howe K.L., Ashurst J.L., Fulton R.S., Sudbrak R., Wen G.,  
RA Jones M.C., Hurler M.E., Andrews T.D., Scott C.E., Searle S.,

RA Ramser J., Whittaker A., Deadman R., Carter N.P., Hunt S.E., Chen R.,  
 RA Cree A., Gunaratne P., Haylak P., Hodgson A., Metzker M.L.,  
 RA Richards S., Scott G., Steffen D., Sodergren E., Wheeler D.A.,  
 RA Worley K.C., Alnsough R., Ambrose K.D., Ansari-Lari M.A., Aradhya S.,  
 RA Ashwell R.I., Babbage A.K., Baggeley C.B., Ballabio A., Banerjee R.,  
 RA Barker G.E., Barlow K.F., Barrett I.P., Bates K.D., Beare D.M.,  
 RA Beasley H., Beasley O., Beck A., Bethel G., Blechschmidt K., Brady N.,  
 RA Bray-Alten S., Bridgeman A.M., Brown A.J., Brown M.J., Bonnin D.,  
 RA Bruford E.A., Buhay C., Burch P., Burford D., Burgess J., Burrill W.,  
 RA Burton J., Bye J.M., Carder C., Carrel L., Chako J., Chapman J.C.,  
 RA Chavez D., Chen E., Chen G., Chen Y., Chen Z., Chinault C.,  
 RA Ciccodicola A., Clark S.Y., Clarke G., Clee C.M., Clegg S.,  
 RA Clerc-Blankenburg K., Clifford K., Cobley V., Cole C.G., Conquer J.S.,  
 RA Corby N., Connor R.E., David R., Davies J., Davis C., Davis J.,  
 RA Delgado O., Deshazo D., Dhami P., Ding Y., Dinh H., Dodsworth S.,  
 RA Draper H., Dugan-Rocha S., Dunham A., Dunn M., Durbin K.J., Dutta I.,  
 RA Eades T., Ellwood M., Emery-Cohen A., Errington H., Evans K.L.,  
 RA Faulkner L., Francis F., Frankland J., Fraser A.E., Galgoczy P.,  
 RA Gilbert J., Gill R., Gloeckner G., Gregory S.G., Gribble S.,  
 RA Griffiths C., Grocock R., Gu Y., Gwilliam R., Hamilton C., Hart E.A.,  
 RA Hawes A., Heath P.D., Heitmann K., Hennig S., Hernandez J.,  
 RA Hjazmann B., Ho S., Hofis M., Howden P.J., Huckle E.J., Hume J.,  
 RA Hut P.J., Hunt A.R., Isherwood J., Jacob L., Johnson D., Jones S.,  
 RA de Jong P.J., Joseph S.S., Keenan S., Kelly S., Kershaw J.K., Khan Z.,  
 RA Kioschis P., Klages S., Knights A.J., Kosiura A., Kovar-Smith C.,  
 RA Laird G.K., Langford C., Lawlor S., Leversha M., Lewis L., Liu W.,  
 RA Lloyd C., Lloyd D.M., Louleghed H., Loveland J.E., Lovell J.D.,  
 RA Lozadó R., Lu J., Lyne R., Ma J., Maheshwari M., Matthews L.H.,  
 RA McDowall J., McLaren S., McMurray A., Meldi P., Meitinger T.,  
 RA Milne S., Miner G., Misty S.L., Morgan M., Morris S., Mueller I.,  
 RA Mullikin J.C., Nguyen N., Nordieck G., Nyakatura G., O'dell C.N.,  
 RA Okwunu G., Palmer S., Pandian R., Parker D., Parrish J.,  
 RA Pasternak S., Patel D., Pearce A.V., Pearson D.M., Pelan S.E.,  
 RA Perez L., Porter K.M., Ramsey J., Reichwald K., Rhodes S.,  
 RA Ridler K.A., Schlössinger D., Schueler M.G., Sehra H.K.,  
 RA Shaw-Smith C., Shen H., Sheridan E.M., Showkeen R., Skuce C.D.,  
 RA Smith M.L., Sotharan E.C., Steingruber H.E., Steward C.A., Storey R.,  
 RA Swann R.M., Swarbreck D., Tabor P.E., Taudien S., Taylor T.,  
 RA Teague B., Thomas K., Thorpe A., Timms K., Tracey A., Trevanion S.,  
 RA Tromans A.C., d'Urso M., Verdusco D., Villaseña D., Waldron L.,  
 RA Wall M., Wang Q., Warren J., Warry G.B., Wei X., West A.,  
 RA Whitehead S.L., Williams M.N., Wilkinson J.E., Willey D.L.,  
 RA Williams G., Williams L., Williamson A., Williamson H., Wilming L.,  
 RA Woodmansey R.L., Wray P.W., Yen J., Zhang J., Zhou J., Zoghbi H.,  
 RA Zorilla S., Buck D., Reinhardt R., Poustka A., Rosenthal A.,  
 RA Lehrach H., Meindl A., Minx P.J., Hillier L.W., Willard H.F.,  
 RA Wilson R.K., Waterston R.H., Rice C.M., Vaudin M., Coulson A.,  
 RA Nelson D.L., Weinstein G., Sulston J.E., Durbin R., Hubbard T.,  
 RA Gibbs R.A., Beck S., Rogers J., Bentley D.R.;  
 RT "The DNA sequence of the human X chromosome.";  
 RL Nature 434:325-337(2005).  
 RN [15]  
 RN TISSUE SPECIFICITY, AND ALTERNATIVE SPLICING.  
 RP Kobiela K., Kobiela A., Trzciak W.H.;  
 RA "Expression of a novel transcript isoform of the EDA gene in human  
 RT umbilical cord.";  
 RT Eur. J. Hum. Genet. Suppl. 7:104-104(1999).  
 RL [16]  
 RN RECEPTOR INTERACTION (ISOFORMS A1 AND A2).  
 RP MEDLINE=20495245; PubMed=1103935; DOI=10.1126/science.290.5491.523;  
 RX Yan M., Wang L.-C., Hymowitz S.G., Schilbach S., Lee J., Goddard A.,  
 RA de Vos A.M., Gao W.-Q., Dixit V.M.;  
 RA "Two-amino acid molecular switch in an epithelial morphogen that  
 RT regulates binding to two distinct receptors.";  
 RL Science 290:523-527(2000).  
 RN [17]  
 RN PROCESSING, MUTAGENESIS OF ARG-153, AND CHARACTERIZATION OF VARIANT  
 RP HIS-156  
 RX MEDLINE=21205766; PubMed=11309369; DOI=10.1093/hmg/10.9.953;  
 RA Elomaa O., Pulkkinen K., Hannellius U., Mikkola M., Saarialho-Kere U.,  
 RA Kere J.;  
 RT "Ectodysplasin is released by proteolytic shedding and binds to the  
 RT EDAR protein.";

RL Hum. Mol. Genet. 10:953-962(2001).  
 RN [8]  
 RN CHARACTERIZATION OF VARIANTS CYS-155; CYS-156 AND HIS-156, MUTAGENESIS  
 RP OF ARG-153; LYS-158 AND ARG-159, AND CLEAVAGE SITE.  
 RX MEDLINE=21309995; PubMed=11416205; DOI=10.1073/pnas.131076098;  
 RX Chen Y., Molloy S.S., Thomas L., Gamble J., Baechinger H.P.,  
 RA Ferguson B.M., Zonana J., Thomas G., Morris N.P.;  
 RA "Mutations within a furin consensus sequence block proteolytic release  
 RT of ectodysplasin-A and cause X-linked hypohidrotic ectodermal  
 RT dysplasia.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7218-7223(2001).  
 RL [9]  
 RN X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 230-391, AND SUBUNIT.  
 RP PubMed=14656435;  
 RX Starowicz S.G., Compaan D.M., Yan M., Wallweber H.J., Dixit V.M.,  
 RA Hwarvasnik M.A., de Vos A.M.;  
 RA "The crystal structures of EDA-A1 and EDA-A2: splice variants with  
 RT distinct receptor specificity.";  
 RL Structure 11:1513-1520(2003).  
 RN [10]  
 RP VARIANT EDA TYR-54.  
 RX MEDLINE=98292028; PubMed=9630076;  
 RA Hertz J.M., Noergaard Hansen K., Juncker I., Kjelsgaard M.,  
 RA Gregersen N.;  
 RA "A novel missense mutation (402C-->T) in exon 1 in the EDA gene in a  
 RT family with X-linked hypohidrotic ectodermal dysplasia.";  
 RL Clin. Genet. 53:205-209(1998).  
 RN [11]  
 RP VARIANT EDA LYS-63.  
 RX MEDLINE=98162231; PubMed=9507389;  
 RA Ferguson B.M., Thomas N.S.T., Munoz F., Morgan D., Clarke A.,  
 RA Zonana J.;  
 RA "Scarcity of mutations detected in families with X linked hypohidrotic  
 RT ectodermal dysplasia: diagnostic implications.";  
 RL J. Med. Genet. 35:112-115(1998).  
 RN [12]  
 RP VARIANT EDA ARG-55.  
 RX MEDLINE=99399307; PubMed=10469321;  
 RX DOI=10.1046/j.1523-1747.1999.00656.x;  
 RA Martinez F., Millan J.M., Orellana C., Prieto F.;  
 RA "X-linked anhidrotic (hypohidrotic) ectodermal dysplasia caused by a  
 RT novel mutation in EDAL gene: 406T > G (Leu55Arg).";  
 RL J. Invest. Dermatol. 113:285-286(1999).  
 RN [13]  
 RP VARIANTS EDA ARG-60; TYR-252; VAL-269; SER-302 AND MET-378.  
 RX MEDLINE=21272350; PubMed=11378824; DOI=10.1038/sj.ejhg.5200635;  
 RA Vincent M.C., Biancalana V., Ginisty D., Mandel J.L., Calvas P.;  
 RA "Mutational spectrum of the EDI gene in X-linked hypohidrotic  
 RT ectodermal dysplasia.";  
 RL Eur. J. Hum. Genet. 9:355-363(2001).  
 RN [14]  
 RP VARIANTS EDA CYS-156; HIS-156; CYS-255; ASP-255; GLY-274; TYR-332 AND  
 RP THR-349.  
 RX MEDLINE=21193173; PubMed=11295832; DOI=10.1002/humu.33;  
 RA Paeskoenen K., Cambiaghi S., Novelli G., Ouzts L.V., Penttinen M.,  
 RA Kere J., Srivastava A.K.;  
 RA "The mutation spectrum of the EDA gene in X-linked anhidrotic  
 RT ectodermal dysplasia.";  
 RL Hum. Mutat. 17:349-349(2001).  
 CC -I- FUNCTION: Seems to be involved in epithelial-mesenchymal signaling  
 CC during morphogenesis of ectodermal organs. Isoform A1 binds only  
 CC to the receptor EDAR, while isoform A2 binds exclusively to the  
 CC receptor XEDAR.  
 CC -I- SUBUNIT: Homotrimer. The homotrimers may then dimerize and form  
 CC higher order oligomers.  
 CC -I- INTERACTION:  
 CC Q9UNEO:EDAR; NExp=1; Intact=EBI-529425, EBI-529289;  
 CC -I- SUBCELLULAR LOCATION: Type II membrane protein and secreted.  
 CC -I- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=8;  
 CC Comment=Additional isoforms seem to exist;  
 CC Name=A1, Synonyms=II;  
 CC IsoId=Q92838-1; Sequence=Displayed;



Qy 361 AVRWVHADISINNSKTTTFFGAIRLGEAPAS 391  
Db 361 AVRWVHADISINNSKTTTFFGAIRLGEAPAS 391

RESULT 3

ED A MOUSE STANDARD; PRT; 391 AA.  
AC OS4693; Q35705; Q9QWJ8; Q9Q201; Q9Q202;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Ectodysplasin A (EDA protein homolog) (Tabby protein) [Contains:  
DE Ectodysplasin A, membrane form; Ectodysplasin A, secreted form].  
GN Name=Eda; Synonyms=Ed1, Ta;  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridae; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
[1]  
RN NUCLEOTIDE SEQUENCE (ISOFORMS TAA; TAB AND TAC).  
RC STRAIN=129/SV;  
RX MEDLINE=96058770; PubMed=9371801; DOI=10.1073/pnas.94.24.13069;  
RA Srivastava A.K., Plespa J., Hartung A.J., Du Y., Ezer S., Jenks T.,  
RA Shimada T., Pekkanen M., Mikkola M.L., Ko M.S.H., Thesleff I.,  
RA Kere J., Schlesinger D.;  
RT "The Tabby phenotype is caused by mutation in a mouse homologue of the  
RT EDA gene that reveals novel mouse and human exons and encodes a  
RT protein (ectodysplasin-A) with collagenous domains."  
RL Proc. Natl. Acad. Sci. U.S.A. 94:13069-13074(1997).  
[2]  
RN NUCLEOTIDE SEQUENCE (ISOFORM TAD).  
RX MEDLINE=97449184; PubMed=9285798; DOI=10.1093/hmg/6.9.1589;  
RA Ferguson B.M., Brockdorff N., Formstone E., Ngyuen T.,  
RA Kronmiller J.E., Zonana J.;  
RT "Cloning of Tabby, the murine homolog of the human EDA gene: evidence  
RT for a membrane-associated protein with a short collagenous domain."  
RL Hum. Mol. Genet. 6:1589-1594(1997).  
[3]  
RN NUCLEOTIDE SEQUENCE (ISOFORMS TA-A2 AND TA-A3).  
RC TISSUE=Embryo;  
RX MEDLINE=20005791; PubMed=10534613; DOI=10.1016/S0925-4773(99)00180-X;  
RA Mikkola M.L., Plespa J., Pekkanen M., Paulin L., Nieminen P., Kere J.,  
RA Thesleff I.;  
RT "Ectodysplasin, a protein required for epithelial morphogenesis, is a  
RT novel TNF homologue and promotes cell-matrix adhesion."  
RL Mech. Dev. 88:133-146(1999).  
CC -1- FUNCTION: Involved in epithelial-mesenchymal signaling during the  
CC morphogenesis of ectodermal organs. Isoform TAA binds only to the  
CC receptor EDAR, while isoform TA-A2 binds exclusively to the  
CC receptor XEDAR.  
CC -1- SUBUNIT: Homotrimer (By similarity).  
CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By  
CC similarity).  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=6;  
CC Comment=Additional isoforms seem to exist;  
CC Name=TAA; Synonyms=Al;  
CC IsoID=OS4693-1; Sequence=VSP\_006471;  
CC Name=TA-A2;  
CC IsoID=OS4693-2; Sequence=VSP\_006471;  
CC Name=TA-A3;  
CC IsoID=OS4693-3; Sequence=VSP\_006469, VSP\_006471;  
CC Name=TAB;  
CC IsoID=OS4693-4; Sequence=VSP\_006466, VSP\_006467;  
CC Name=TAC;  
CC IsoID=OS4693-5; Sequence=VSP\_006465, VSP\_006469;  
CC Name=TAD;  
CC IsoID=OS4693-6; Sequence=VSP\_006470;  
CC -1- PTM: N-glycosylated (By similarity).  
CC Comment=Secreted form (By similarity).







ID	Q4SJR1_TETNG PRELIMINARY; PRT; 207 AA.
AC	Q4SJR1;
DT	13-SEP-2005 (TrEMBLrel. 31, Created)
DD	13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE	13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE	Chromosome 1 SCAP14573, whole genome shotgun sequence. (fragment).
GN	ORFNames=GSTENG00017090001;
OS	Tetraodon nigroviridis (Green puffer).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC	Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC	Tetraodontoidea; Tetraodontidae; Tetraodon.
NCBI_TaxID=99883;	
RN	[1]
RP	NUCLEOTIDE SEQUENCE.
RA	Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N., Mauceli B., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A., Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B., Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S., Anthonard V., Jubin C., Castelli V., Katinka M., Vacherie B., Blement C., Skalli Z., Cattolico L., Poullain J., De Berardinis V., Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzou J., Parra G., Lardier G., Chappelle C., McKernan K.J., McEwan P., Bosak S., Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J., Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M., Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C., Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.; "Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype."; Nature 431:946-957(2004).
RN	[2]
RP	NUCLEOTIDE SEQUENCE.
RG	Genoscope; Whitehead Institute Centre for Genome Research;
RL	Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC	-1- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is preliminary data.
CC	EMBL; CAEA01014573; CAF99121.1; -; Genomic_DNA.
DR	InterPro; IPR008160; Collagen.
DR	InterPro; IPR006052; TNF_family.
DR	Pfam; PF01391; Collagen; 1.
DR	Pfam; PF00229; TNF; 1.
DR	PROSITE; PS00251; TNF_1; UNKNOWN_1.
DR	PROSITE; PS50049; TNF_2; 1.
KW	Collagen.
FT	NON_TER . 1 1
FT	NON_TER 207 207
SQ	SEQUENCE 207 AA; 22149 MW; 6CB874AF177DA5CA CRC64;
Query Match	38.6%; Score 809; DB 2; Length 207;
Best Local Similarity	68.9%; Pred. No. 6e+40;
Matches 142; Conservative	32; Mismatches 32; Indels 0; Gaps 0;
Qy	186 GPFGPPGPGPQGPGIPGIPGTITVWGPPGPPGPGLQSPGGAADKAGTRE 245
Dd	1 GPFGPPGPGPQGPGIPGIPGSGNVGVGPPGPPGPGPTQTGGPDPKTKTRE 60
Qy	246 NPQAVHVLGGCGSAIQVNDSLGSGLVNDWSRITWNPKVKLHPRSGEVLVDGYTFIYS 305
Dd	61 FORAVHVLGGQETTIOVRDLSEGLIRNMKNVISIHHRVFNKHSRSGEVLVDGYTFIYS 120
Qy	306 QVEYVYNFTDFASYEVVWDKPFLQCTRSJETGKNTNTCYTAGVCLLKARQKIADVW 365
Dd	121 QVEYVYNFTDIASYEVVWDSNPFLRCTSIETQRKENTCYTAGVSLLRAGRISRIIV 180
Qy	366 HADISINMSKHHTFFGAIRLGEAPAS 391
Dd	181 YEDTLISMTNHTTFLGSRVLEAPSA 206

RESULT 9  
Q8VIF6 MOUSE

ID	QB51P6	MOUSE PRELIMINARY;	PRT;	742 AA.
AC	QB51P6;			
DT	01-MAR-2002	(TrEMBLrel. 20, Created)		
DT	01-MAR-2002	(TrEMBLrel. 20, Last sequence update)		
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)		
DT	Scavenger receptor with C-type lectin.			
GN	Scavenger receptor with C-type lectin.			
GN	Names=Colect12; Synonyms=srcl1;			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;			
OC	Muridae; Murinae; Mus.			
OC	NCBI_TaxID=10090;			
OC	[1]			
RN	NCU012515			
RP	Medline=21575632; PubMed=11718900; DOI=10.1016/S0167-4781(01)00284-6;			
RP	Medline=21575632; PubMed=11718900; DOI=10.1016/S0167-4781(01)00284-6;			
RT	Nakamura K., Funakoshi H., Tokunaga F., Nakamura T.;			
RT	"Molecular cloning of a mouse scavenger receptor with C-type lectin			
RT	(SRC1) (1), a novel member of the scavenger receptor family.";			
RL	Biochim. Biophys. Acta 1522:53-58(2001).			
RL	EMBL; AB038519; BAB2497.1; -; mRNA.			
DR	HSSP; P22897; 1EGI.			
DR	MGI; MGI:2152907; Colect12.			
DR	GO; GO:0016021; C:integral to membrane; TAS.			
DR	GO; GO:0016020; C:membrane; ISS.			
DR	GO; GO:0005534; F:galactose binding; ISS.			
DR	GO; GO:0030169; F:low-density lipoprotein binding; ISS.			
DR	GO; GO:0008329; F:pattern recognition receptor activity; ISS.			
DR	GO; GO:0005044; F:scavenger receptor activity; ISS.			
DR	GO; GO:0019733; P:antibacterial humoral response (sensu Verte. .; ISS.			
DR	GO; GO:0019734; P:antifungal humoral response (sensu Vertebrata); ISS.			
DR	GO; GO:0005056; P:carbohydrate mediated signaling; ISS.			
DR	GO; GO:0045087; P:innate immune response; ISS.			
DR	GO; GO:0045194; P:oxidized low-density lipoprotein catabolism; ISS.			
DR	GO; GO:0006910; P:phagocytosis, recognition; ISS.			
DR	GO; GO:0051260; P:protein homooligomerization; ISS.			
DR	InterPro; IPR002353; AntifreezeZell.			
DR	InterPro; IPR008160; Collagen.			
DR	InterPro; IPR001304; Lectin_C.			
DR	Pfam; PF01391; Collagen; 2.			
DR	Pfam; PF00059; Lectin_C; 1.			
DR	PRINTS; PR00356; ANTI-FREEZEZELL.			
DR	SMART; SM00034; CLECT; 1.			
DR	PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.			
DR	PROSITE; PS00041; C_TYPE_LLECTIN_2; 1.			
KW	Lectin; Receptor.			
SEQ	SEQUENCE 742 AA; 81306 MW; 85A90D3AE881DB6B CRC64;			
	Query Match	13.6%; Score 286; DB 2; Length 742;		
	Best Local Similarity	30.5%; Pred. No. 9.2e-09;		
	Matches	84; Conservative 29; Mismatches 96; Indels 66; Gaps 11;		
Qy	82	GTPTGTSLLSLGLDLPDPSITSHLGQSPKQQLPFGAALHSDSDGQHMALNNFFP 141		
Db	443	GPPEGRGPKDGRSGQPGP-TGNKQKGKGRGPGPGA-----GEGTIGPVGP 493		
Qy	142	DEKPYSEESRRVRNRKRSKSNAGDPVKNKKKGKAGPGPGNGPPGPP 195		
Db	494	GE-----RGSKSGSGQPKSGRSGPKPGQPGSGDPGPPGPKDGLPG 539		
Qy	196	PQGPPIPGI-----PGIGTT-VMGPPGPPGPPGPPGPGPGQPGSGADKAGTRENOP 248		
Db	540	PQGPPIPGI-----PGIGTT-VMGPPGPPGPPGPPGPGPGQPGSGADKAGTRENOP 248		
Qy	249	AVVHLQGSAGIQVNDLSGVLNDSRIWNPVKLHPRSGLEVLVDGTIFYISQVE 308		
Db	599	T-----PASEV-----NGCPHHWNFTDKCYFSL-----EKEILEDKLEFC----- 635		
Qy	309	VYVINFDFASVEVWDEKPFQCTRSIETGKTY 343		
Db	636	-----EDKSHLVFINSREEQWIKKHTVGRSH 664		
RESULT	10			







DR PROSITE; PS50041; C\_TYPE\_LECTIN\_2; 1.  
SQ SEQUENCE 742 AA; 81289 MW; I350DCCB34522C45 CRC64;

Query Match 13.4%; Score 282; DB 2; Length 742;  
Best Local Similarity 30.2%; Pred. No. 1.6e-08;  
Matches 83; Conservative 29; Mismatches 97; Indels 66; Gaps 11;

QY 82 GTPGTSGLTSSLGGLDPSPTSHLQGPSKPOKPLEGAEALHSQDQHALLNFFPP 141  
DB 443 GPGPGPKGDRSGSQPPGP-TGNKGQKGEKEGPPGPA-----GERGTIGVGP 493  
QY 142 DERPYSEESRRVRRNKRKSKNEGADGVKNKKKGKAGPPGPGNPPGPPG 195  
DB 494 GE-----RSGKSGSGKSGRSGPKPGQPSGDPGPPGPKDGLPG 539  
QY 196 PQGPPGPGI-----PGIPGTT-VMGPPGPPGPPGQPSGAGADKAGTRENQP 248  
DB 540 PQGPPGPGIQTGTVGFGVGRGLPGLPVCPMPGPKGPPGPPGSGAMEPLAL-QNEP 598  
QY 249 AVVHQQGSGATQVKNLDSGLVNDWSRLTMNPKVKLHPRSGLEVLVDGTYFYISQVE 308  
DB 599 T-----PASEV-----NGCPHPKNTDKCYFSL-----EKEIFEDAKLFC----- 635  
QY 309 VVYINFDFASVYVVVDKPFLLQCTRSIETGKTY 343  
DB 636 -----EDKSSHLVFNREEQWIKHTVGRSH 664

RESULT 13  
Q59EE7 HUMAN  
ID Q59EE7 HUMAN PRELIMINARY; PRT; 1792 AA.  
AC Q59EE7;  
DT 10-MAY-2005 (TrEMBLrel. 30, Created)  
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)  
DE 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)  
DB Pro-alpha-1 type V collagen variant (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Aorta endothelial cell;  
RA Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,  
RA Ohara O., Nagase T., Kikuno F.R.;  
RT "None Title";  
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB209864; BAD93101.1; -; mRNA.  
DR GO; GO:0005581; C:collagen; IEA.  
DR GO; GO:0005737; C:cytoplasm; IEA.  
DR GO; GO:0005201; F:extracellular matrix structural constituent; IEA.  
DR GO; GO:0007155; P:cell adhesion; IEA.  
DR GO; GO:0006817; P:phosphate transport; IEA.  
DR InterPro; IPR008160; Collagen.  
DR InterPro; IPR00885; Fib\_collagen\_C.  
DR InterPro; IPR001791; Laminin G.  
DR InterPro; IPR003129; Laminin\_G\_TSP\_N.  
DR Pfam; PF01410; COLFI; 1.  
DR SMART; SM00391; Collagen; 18.  
DR SMART; SM0038; COLFI; 1.  
DR SMART; SM00282; LamG; 1.  
DR SMART; SM00210; TSPN; 1.  
DR KEGG; K00001; Collagen; Extracellular matrix.  
FT NON TER 1  
SQ SEQUENCE 1792 AA; 178517 MW; E47B3093C34C5A9D CRC64;

Query Match 13.4%; Score 280.5; DB 2; Length 1792;  
Best Local Similarity 30.1%; Pred. No. 4.7e-08;  
Matches 83; Conservative 9; Mismatches 69; Indels 115; Gaps 10;

QY 13 PAAAPRBRGSGC-GCGGAPARAGEGNSCLLFLGFGLSLALHLTLCCVLELRSELRR 71

RP DISEASE, AND VARIANT EDS-I SER-1639.  
RX MEDLINE=97195540; PubMed=9042913;  
RA de Faeppe A., Nuytink L., Hausser I., Anton-Lamprecht I.,  
RA Naeyaert J.-M.,  
RT "Mutations in the COL5A1 gene are causal in the Ehlers-Danlos  
RT syndromes I and II";  
RL Am. J. Hum. Genet. 60:547-554(1997).  
RN [7]  
RP VARIANTS EDS-1 SER-530 AND ASP-1489.  
RX MEDLINE=2068401; PubMed=10602121;  
RX DOI=10.1002/(SICI)1096-8628(200010)30:1<72::AID-AJMG13>3.0.CO;2-C;  
RA Giunta C., Steinmann B.,  
RT "Compound heterozygosity for a disease-causing G1489D and disease-  
RT modifying G530S substitution in COL5A1 of a patient with the classical  
RT type of Ehlers-Danlos syndrome: an explanation of intrafamilial  
RT variability?";  
RL Am. J. Med. Genet. 90:72-79(2000).  
CC -!- FUNCTION: Type V collagen is a member of group I collagen  
CC (fibrillar forming collagen). It is a minor connective tissue  
CC component of nearly ubiquitous distribution. Type V collagen binds  
CC to DNA, heparan sulfate, thrombospondin, heparin, and insulin.  
CC -!- SUBUNIT: Trimers of two alpha 1(V) and one alpha 2(V) chains in  
CC most tissues and trimers of one alpha 1(V), one alpha 2(V), and  
CC one alpha 3(V) chains in placenta.  
CC -!- PTM: Prolines at the third position of the tripeptide repeating  
CC unit (G-X-Y) are hydroxylated in some or all of the chains.  
CC -!- PTM: Sulfated on 40% of tyrosines.  
CC -!- DISEASE: Defects in COL5A1 are a cause of Ehlers-Danlos syndrome  
CC type I (EDS-I) [MIM:130000]; also known as Ehlers-Danlos syndrome  
CC gravis. EDS-I is a connective-tissue disorder characterized by  
CC loose-jointedness and fragile, velvety, stretchable, bruisable  
CC skin that heals with peculiar 'cigarette-paper' scars. Inheritance  
CC is autosomal dominant.  
CC -!- DISEASE: Defects in COL5A1 are a cause of Ehlers-Danlos syndrome  
CC type II (EDS-II) [MIM:130010]; also known as Ehlers-Danlos  
CC syndrome mitis. Inheritance is autosomal dominant.  
CC -!- SIMILARITY: Belongs to the fibrillar collagen family.  
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
DR EMBL; D90279; BA014323.1; -, mRNA.  
DR PDB; 1A89; Model; A/B/C=904-924.  
DR PDB; 1A9A; Model; A/C=904-924.  
DR Ensembl; ENSG00000130635; Homo sapiens.  
DR HGNC; HGNC:2209; COL5A1.  
DR MIM; 120215; -.  
DR MIM; 130000; -.  
DR MIM; 130010; -.  
DR GO; GO:0005588; C:collagen type V; TAS.  
DR InterPro; IPR008161; Clq helix.  
DR InterPro; IPR008160; Collagen.  
DR InterPro; IPR000885; Fib\_collagen\_C.  
DR InterPro; IPR01791; Laminin G.  
DR InterPro; IPR003129; Laminin G\_TSP\_N.  
DR Pfam; PF01410; COLFI; 1.  
DR Pfam; PF01391; Collagen; 18.  
DR ProDom; PD000007; Clq\_helix; 4.  
DR ProDom; PD002078; Fib\_collagen\_C; 1.  
DR SMART; SM00038; COLFI; 1.  
DR SMART; SM00282; LamG; 1.  
DR SMART; SM00210; TSPN; 1.  
KW 3D-structure; Collagen; Direct cellular sequencing; Disease mutation;  
KW Ehlers-Danlos syndrome; Extracellular matrix; Heparin-binding;  
KW Hydroxylation; Repeat; Signal; Structural protein; Sulfation.  
FT SIGNAL 1 37 Potential.  
FT CHAIN 38 1605 Collagen alpha 1(V) chain.  
FT PROPEP 1606 1838 C-terminal propeptide.  
FT DOMAIN 39 230 TSP N-terminal.

FT REGION 231 443 Nonhelical region.  
FT REGION 444 558 Interrupted collagenous region.  
FT REGION 559 1570 Triple-helical region.  
FT REGION 1571 1605 Nonhelical region.  
FT MOD\_RES 234 234 Sulfotyrosine (Potential).  
FT MOD\_RES 236 236 Sulfotyrosine (Potential).  
FT MOD\_RES 240 240 Sulfotyrosine (Potential).  
FT MOD\_RES 262 262 Sulfotyrosine (Potential).  
FT MOD\_RES 263 263 Sulfotyrosine (Potential).  
FT MOD\_RES 338 338 Sulfotyrosine (Potential).  
FT MOD\_RES 340 340 Sulfotyrosine (Potential).  
FT MOD\_RES 346 346 Sulfotyrosine (Potential).  
FT MOD\_RES 347 347 Sulfotyrosine (Potential).  
FT MOD\_RES 416 416 Sulfotyrosine (Potential).  
FT MOD\_RES 417 417 Sulfotyrosine (Potential).  
FT MOD\_RES 420 420 Sulfotyrosine (Potential).  
FT MOD\_RES 421 421 Sulfotyrosine (Potential).  
FT MOD\_RES 570 570 Hydroxyproline.  
FT MOD\_RES 576 576 Hydroxyproline.  
FT MOD\_RES 621 621 Hydroxyproline.  
FT MOD\_RES 627 627 5-hydroxylysine.  
FT MOD\_RES 639 639 Hydroxyproline.  
FT MOD\_RES 642 642 5-hydroxylysine.  
FT MOD\_RES 648 648 Hydroxyproline.  
FT MOD\_RES 654 654 Hydroxyproline.  
FT MOD\_RES 657 657 Hydroxyproline.  
FT MOD\_RES 675 675 Hydroxyproline.  
FT MOD\_RES 678 678 Hydroxyproline.  
FT MOD\_RES 680 680 Hydroxyproline.  
FT MOD\_RES 686 686 Hydroxyproline.  
FT MOD\_RES 690 690 Hydroxyproline.  
FT MOD\_RES 696 696 Hydroxyproline.  
FT MOD\_RES 705 705 Hydroxyproline.  
FT MOD\_RES 708 708 5-hydroxylysine.  
FT MOD\_RES 717 717 Hydroxyproline.  
FT MOD\_RES 720 720 Hydroxyproline.  
FT MOD\_RES 726 726 Hydroxyproline.  
FT MOD\_RES 732 732 Hydroxyproline.  
FT MOD\_RES 744 744 5-hydroxylysine.  
FT MOD\_RES 750 750 Hydroxyproline.  
FT MOD\_RES 756 756 Hydroxyproline.  
FT MOD\_RES 762 762 Hydroxyproline.  
FT MOD\_RES 765 765 Hydroxyproline.  
FT MOD\_RES 771 771 Hydroxyproline.  
FT MOD\_RES 774 774 5-hydroxylysine.  
FT MOD\_RES 780 780 Hydroxyproline.  
FT MOD\_RES 789 789 Hydroxyproline.  
FT MOD\_RES 795 795 5-hydroxylysine.  
FT MOD\_RES 804 804 5-hydroxylysine.  
FT MOD\_RES 807 807 5-hydroxylysine.  
FT MOD\_RES 810 810 Hydroxyproline.  
FT MOD\_RES 816 816 Hydroxyproline.  
FT MOD\_RES 819 819 5-hydroxylysine.  
FT MOD\_RES 834 834 Hydroxyproline.  
FT MOD\_RES 846 846 5-hydroxylysine.  
FT MOD\_RES 861 861 Hydroxyproline.  
FT MOD\_RES 864 864 5-hydroxylysine.  
FT MOD\_RES 870 870 Hydroxyproline.  
FT MOD\_RES 873 873 Hydroxyproline.  
FT MOD\_RES 876 876 Hydroxyproline.  
FT MOD\_RES 882 882 5-hydroxylysine.  
FT MOD\_RES 888 888 Hydroxyproline.  
FT MOD\_RES 891 891 Hydroxyproline.  
FT MOD\_RES 897 897 5-hydroxylysine.  
FT MOD\_RES 903 903 Hydroxyproline.  
FT MOD\_RES 906 906 Hydroxyproline.  
FT MOD\_RES 930 930 Hydroxyproline.  
FT MOD\_RES 945 945 Hydroxyproline.  
FT MOD\_RES 1017 1017 Hydroxyproline.  
FT MOD\_RES 1020 1020 Hydroxyproline.  
FT MOD\_RES 1023 1023 Hydroxyproline.  
FT MOD\_RES 1029 1029 Hydroxyproline.  
FT MOD\_RES 1221 1221 Hydroxyproline.







Db	1	MGYPEVRRBPLPAAAAPRRSGSCGCRGAPARAGEGNSCTLFTLFGFLSLALHLLTCC	60
Qy	61	YLRLSLRLRRGNAESRLGSGTGTGTTSSLGGLDPDPSITSHLGQSPKQOPLRGE	120
Db	61	YLRLSLRLRRERGTTESLGLGPGAGTGTGTTSSGSLDPVPGITRHLGQSPQOPLRGE	120
Qy	121	AAHLSDSQDGHOMALLNFFPDDEKPYSEESRRYVRNKRKSKSNEGADGPVNKKKKKKAK	180
Db	121	DPUPPESQDRHOMALLNFFPDDEKAYSEESRRYVRNKRKSKSGGADGPVNKKKKKKAK	180
Qy	181	PPGPNGGPPGPPGPPGPGIPGIPGTGTVMGPPGPPGPPGPPGPGGPGSGAADK	240
Db	181	PPGPNGGPPGPPGPPGPPGPGIPGIPGTGTVMGPPGPPGPPGPPGPGGPGSGAADK	240
Qy	241	AGTRENQPAVVHLQGQSGSAIQVKNDSLGGVLNDWSRITMNPVKFLHPRSGEVLVDGT	300
Db	241	TGTRENQPAVVHLQGQSGSAIQVKNDSLGGVLNDWSRITMNPVKFLHPRSGEL	293
Qy	301	YFIYSQVEVYYINFTDFASVEVVVDKXPFLQCTRSIETGKTNYNCTYTAGVCLLKAROKI	360
Db	294	-----EVYYINFTDFASVEVVVDKXPFLQCTRSIETGKTNYNCTYTAGVCLLKAROKI	346
Qy	361	AVKMVHADISINMSKHTFFGAIIRLGPAPAS	391
Db	347	AVKMVHADISINMSKHTFFGAIIRLGPAPAS	377

```

RESULT 5
US-09-949-016-11119
; Sequence 11119, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11119
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11119

```

```

Qy 372 NMSKHTTFFGAIRLGEAPAS 391
    ||||||||||||||||||
Db 241 NMSKHTTFFGAIRLGEAPAS 260

RESULT 6
US-07-609-716-66
; Sequence 66, Application US/07609716
; Patent No. 5514581
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: Functional Recombinantly Prepared
; TITLE OF INVENTION: Synthetic Protein Polymer
; NUMBER OF SEQUENCES: 118
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/609,716
; FILING DATE: 06-NOV-1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: A-55186-3/BIR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 66:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 357 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-609-716-66

Query Match 14.0%; Score 294; DB 1; Length 357;
Best Local Similarity 32.4%; Pred. No. 1.1e-16;
Matches 69; Conservative 9; Mismatches 71; Indels 64; Gaps 3;

Qy 71 ERGAESRLGSGTGPTSGTSLSSLGGLDPSF--ITSHLGQPSPKQQLPEGCEAALHSDSQ 128
Db 71 DRGDAGPKGADGSPGAPGVSGAGPCGPPGPPGPPGAPGPPGPPGPPGPPGPPGPPG 130

Qy 129 DGHQWALLNFFPDPEKPYSEESRRVERNKRSKNEGADGPFVNKKKKKAGKAGP 188
Db 131 D-----RGDAGPKGADGSPGAPGVSGAGPPGPPGPPGPP 165

Qy 189 GPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 212
Db 166 GPPGAPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 225

Qy 213 -VMGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 244
Db 226 GPPGAPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 258

RESULT 7
US-08-642-255-33
; Sequence 33, Application US/08642255
; Patent No. 5773249
; GENERAL INFORMATION:

```

```

; APPLICANT: CAPPELLO, Joseph
; APPLICANT: FERRARI, Franco A.
; TITLE OF INVENTION: High Molecular Weight Collagen-Like
; TITLE OF INVENTION: Protein Polymers
; NUMBER OF SEQUENCES: 135
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/642,255
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: ROWLAND, Berttram I.
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A55556-3/BIR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 357 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-642-255-33

Query Match 14.0%; Score 294; DB 1; Length 357;
Best Local Similarity 32.4%; Pred. No. 1.1e-16;
Matches 69; Conservative 9; Mismatches 71; Indels 64; Gaps 3;

QY 71 ERGAESRLGGSGTGTGTLSSLGGLDPSD--ITSHLGQSPKQOPLPGEAALHSDSQ 128
Db 71 DRGDAGPKGADGSPGAGVSGAGPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 130
QY 129 DGHQWALLNFFPDEKPYSEESRRVRNRKSKSNEGADGVPVKNKKKGKAGPPGPPGPP 188
Db 131 D-----RGDAGPKGADGSPGAGVSGAGPGPPGPPGPPGPPGPPGPPGPPGPP 165
QY 189 GPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 212
Db 166 GPPGAGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 225
QY 213 -VMGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 244
Db 226 GPPGAGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 258

RESULT 8
US-08-475-411A-66
; Sequence 66, Application US/08475411A
; Patent No. 6140072
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: Functional Recombinantly Prepared
; TITLE OF INVENTION: Synthetic Protein Polymer
; NUMBER OF SEQUENCES: 119
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco

; APPLICANT: CAPPELLO, Joseph
; APPLICANT: FERRARI, Franco A.
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,411A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/609,716
; FILING DATE: 06-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/269,429
; FILING DATE: 09-NOV-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/114,618
; FILING DATE: 29-OCT-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/927,258
; FILING DATE: 04-NOV-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-55186-9/RFT/MTK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 66:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 357 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-475-411A-66

Query Match 14.0%; Score 294; DB 2; Length 357;
Best Local Similarity 32.4%; Pred. No. 1.1e-16;
Matches 69; Conservative 9; Mismatches 71; Indels 64; Gaps 3;

QY 71 ERGAESRLGGSGTGTGTLSSLGGLDPSD--ITSHLGQSPKQOPLPGEAALHSDSQ 128
Db 71 DRGDAGPKGADGSPGAGVSGAGPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 130
QY 129 DGHQWALLNFFPDEKPYSEESRRVRNRKSKSNEGADGVPVKNKKKGKAGPPGPPGPP 188
Db 131 D-----RGDAGPKGADGSPGAGVSGAGPGPPGPPGPPGPPGPPGPPGPPGPP 165
QY 189 GPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 212
Db 166 GPPGAGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 225
QY 213 -VMGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 244
Db 226 GPPGAGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 258

RESULT 9
US-08-478-029A-66
; Sequence 66, Application US/08478029A
; Patent No. 6184348
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: Functional Recombinantly Prepared
; TITLE OF INVENTION: Synthetic Protein Polymer
; NUMBER OF SEQUENCES: 119
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
```





```
; APPLICATION NUMBER: US/07/609,716
; FILING DATE: 06-NOV-1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: A-55186-3/BIR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 408 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-609-716-65

Query Match 13.2%; Score 277; DB 1; Length 408;
Best Local Similarity 39.4%; Pred. No. 3.5e-15;
Matches 65; Conservative 5; Mismatches 61; Indels 34; Gaps 4;

QY 79 GCGTPTGTSGLSLGGLDPSDPTSHLQGPSKQPLEPGEAALHSDSQDCHQWALLNF 138
Db 58 GPPGPPGAGPVGSPGAGPPGP-----PGPPGPGAGPPG----- 94
QY 139 FFPDEKPYSEESRRVRNKRKSKSNEGADGPKVKKKKGKAGPPGPNPGPPGPPGQ 198
Db 95 --PPGPPGPPGAGPV-----GSPGAGPPGPPGPPGPPGPPGPPGPPGAG 145
QY 199 PGIGIGIGIPGTTVMGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 243
Db 146 PVGSPGAGPPGPP--GPPGPPGAGPPGPPGPPGPPGPPGPPGPPGPPGPPG 188

RESULT 12
US-08-475-411A-65
; Sequence 65, Application US/08475411A
; Patent No. 6140072
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: Functional Recombinantly Prepared
; TITLE OF INVENTION: Synthetic Protein Polymer
; NUMBER OF SEQUENCES: 119
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,411A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/609,716
; FILING DATE: 06-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/269,429
; FILING DATE: 09-NOV-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/114,618
; FILING DATE: 29-OCT-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/927,258

; APPLICATION NUMBER: US 06/927,258
; FILING DATE: 04-NOV-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-55186-9/RPT/MTK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 408 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-475-411A-65

Query Match 13.2%; Score 277; DB 2; Length 408;
Best Local Similarity 39.4%; Pred. No. 3.5e-15;
Matches 65; Conservative 5; Mismatches 61; Indels 34; Gaps 4;

QY 79 GCGTPTGTSGLSLGGLDPSDPTSHLQGPSKQPLEPGEAALHSDSQDCHQWALLNF 138
Db 58 GPPGPPGAGPVGSPGAGPPGP-----PGPPGPPGAGPPG----- 94
QY 139 FFPDEKPYSEESRRVRNKRKSKSNEGADGPKVKKKKGKAGPPGPNPGPPGPPGQ 198
Db 95 --PPGPPGPPGAGPV-----GSPGAGPPGPPGPPGPPGPPGPPGPPGAG 145
QY 199 PGIGIGIGIPGTTVMGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 243
Db 146 PVGSPGAGPPGPP--GPPGPPGAGPPGPPGPPGPPGPPGPPGPPGPPGPPG 188

RESULT 13
US-08-478-029A-65
; Sequence 65, Application US/08478029A
; Patent No. 6184348
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: Functional Recombinantly Prepared
; TITLE OF INVENTION: Synthetic Protein Polymer
; NUMBER OF SEQUENCES: 119
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,029A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/609,716
; FILING DATE: 06-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/269,429
; FILING DATE: 09-NOV-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/114,618
; FILING DATE: 29-OCT-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/927,258
; FILING DATE: 04-NOV-1986
; ATTORNEY/AGENT INFORMATION:
```

NAME: Trecartin, Richard F.  
REGISTRATION NUMBER: 31,801  
REFERENCE/DOCKET NUMBER: A-55186-8/RPT/MTK  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-781-1989  
TELEFAX: 415-398-3249  
INFORMATION FOR SEQ ID NO: 65:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 408 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-478-029A-65

Query Match 13.2%; Score 277; DB 2; Length 408;  
Best Local Similarity 39.4%; Pred. No. 3.5e-15;  
Matches 65; Conservative 5; Mismatches 61; Indels 34; Gaps 4;

QY 79 GSGTGTGTSGLSLGLDPSPTITSHLQSPKQKQPLEPGEAALHSDSDQDGHQALLNPF 138  
DB 58 GPGGPPGAPGVSGAPGPPG-----PGPPGAPGPPG----- 94  
QY 139 FPPDEKPYSEBSRRVRNKRKSKNEGADGPVNKKKKKAGPPGPPGPPGPPGPPGPPG 198  
DB 95 --PPGPPGPPGAPV-----GSPGAPGPPGPPGPPGPPGPPGPPGPPGPPG 145  
QY 199 PGPIGPIGPIGTTVMGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 243  
DB 146 PVGSGAPGPPGPP--GPPGPPGAPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 188

RESULT 14  
US-08-931-820-1  
; Sequence 1, Application US/08931820  
; Patent No. 6010863  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Assay for collagen degradation  
; NUMBER OF SEQUENCES: 4  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/931,820  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP 96202596.1  
; FILING DATE:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1057 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; TISSUE TYPE: Collagen type I  
US-08-931-820-1

Query Match 12.7%; Score 265.5; DB 2; Length 1057;  
Best Local Similarity 29.7%; Pred. No. 1.1e-13;  
Matches 78; Conservative 16; Mismatches 80; Indels 89; Gaps 9;

QY 13 PAAAPRRGSGQ-----CGCGGAPARAGE-----GNSCLLFLGFFGLSLALHLLTLCCY 61  
DB 613 PAGAPGDKSGSGPPAGPTGARGAPGDRGPPGPA-----GFAG----- 654

QY 62 LELRSELRERGAESRLGGSGTGTGTSGLSLGLDPSPTITSHLQSPKQKQPLEPGEA 121  
DB 655 -----PFGADGQPGKAGSGPDAGAKGADAGPPGAPG-----AGPPGPIGNVGPAG- 700  
QY 122 ALHSDSDQHOMALLNFFPPDEKPYSEBSRRVRNKRKSKNEGADGPVNKKKKKAGKAGP 181  
DB 701 -----KGARSGAGPPGATGPPGAAGRVGP 724  
QY 182 PGPNGPPGPPGPPG-----QGPPGPIGPIGTTVMGPPGPPGPPGPPGPPGPPGPPG 235  
DB 725 PPSGNAGPPGPPGPPGAGKEGKGPGETGPAGPGE--VGPPGPPGAPGKSGPGADGPA 782  
QY 236 GAADKAGTR--ENQPAVVHQQ 256  
DB 783 GARGTGPQGIAGRGVVGLPQ 805

RESULT 15  
US-10-153-469A-16  
; Sequence 16, Application US/10153469A  
; Patent No. 6927287  
; GENERAL INFORMATION:  
; APPLICANT: GRUSKIN, ELLIOT A.  
; BROKAW, JANE  
; ZHANG, GUANGHUI  
; PAOLELLA, DAVID  
; TITLE OF INVENTION: AMINO ACID MODIFIED POLYPEPTIDES  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DILWORTH & BARRESE  
; STREET: 333 EARLE OVINGTON BOULEVARD  
; CITY: UNIONDALE  
; STATE: NY  
; COUNTRY: U.S.A.  
; ZIP: 11553  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/153,469A  
; FILING DATE: 22-May-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/169,768  
; FILING DATE: 09-OCT-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: STEEN, JEFFREY S  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 228-8484  
; TELEFAX: (516) 228-8516  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1057 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:  
US-10-153-469A-16

Query Match 12.7%; Score 265.5; DB 2; Length 1057;  
Best Local Similarity 29.7%; Pred. No. 1.1e-13;  
Matches 78; Conservative 16; Mismatches 80; Indels 89; Gaps 9;

QY 13 PAAAPRRGSGQ-----CGCGGAPARAGE-----GNSCLLFLGFFGLSLALHLLTLCCY 61  
DB 613 PAGAPGDKSGSGPPAGPTGARGAPGDRGPPGPA-----GFAG----- 654  
QY 62 LELRSELRERGAESRLGGSGTGTGTSGLSLGLDPSPTITSHLQSPKQKQPLEPGEA 121





[illegible]



```
Db 241 AGTRENQAVVHLQGGSAIQVKNLDSGGVLDNSRITMNPVKFKLHPRSGEVLVDGT 300
Qy 301 YFIYSQVEVYINFTDFASYEVVVDKPFLOCTRSIETGKTNNTCYTAGVCLLKAROKI 360
Db 301 YFIYSQVEVYINFTDFASYEVVVDKPFLOCTRSIETGKTNNTCYTAGVCLLKAROKI 360
Qy 361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391
Db 361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391

RESULT 7
US-10-719-205-6
; Sequence 6, Application US/10719205
; Publication No. US20040086971A1
; GENERAL INFORMATION:
; APPLICANT: Haldeman, Betty A.
; APPLICANT: Thayer, Edward C.
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: ADIPOCYTE COMPLEMENT RELATED PROTEIN
; FILE REFERENCE: ZACRP3X2
; CURRENT APPLICATION NUMBER: US/10/719,205
; CURRENT FILING DATE: 2003-11-21
; PRIOR FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US/10/012,605C
; PRIOR FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: US 60/254,019
; PRIOR FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-719-205-6

Query Match 100.0%; Score 2098; DB 4; Length 391;
Best Local Similarity 100.0%; Pred. No. 2.6e-137;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGYPEVERRELLPAAAPRERGSGCGCGGAPARAGEGNSCLLFLGFFGLSLALHLLTLCC 60
Db 1 MGYPEVERRELLPAAAPRERGSGCGCGGAPARAGEGNSCLLFLGFFGLSLALHLLTLCC 60
Qy 61 YLELSELRRRGAESRLGGSGTPTGTSLSLGLDPSDPTITSHLGQSPKQPLEPGE 120
Db 61 YLELSELRRRGAESRLGGSGTPTGTSLSLGLDPSDPTITSHLGQSPKQPLEPGE 120
Qy 121 AALHSDSQDGHQMALNFFPDDEKPYSEESRRVRNRKRSKNEGADGVPVKNKKKKKAG 180
Db 121 AALHSDSQDGHQMALNFFPDDEKPYSEESRRVRNRKRSKNEGADGVPVKNKKKKKAG 180
Qy 181 PPGNGPPGPPGPPGPPGPGIPIGPIGTVMGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
Db 181 PPGNGPPGPPGPPGPPGPPGPGIPIGPIGTVMGPPGPPGPPGPPGPPGPPGPPGPPG 240
Qy 241 AGTRENQAVVHLQGGSAIQVKNLDSGGVLDNSRITMNPVKFKLHPRSGEVLVDGT 300
Db 241 AGTRENQAVVHLQGGSAIQVKNLDSGGVLDNSRITMNPVKFKLHPRSGEVLVDGT 300
Qy 301 YFIYSQVEVYINFTDFASYEVVVDKPFLOCTRSIETGKTNNTCYTAGVCLLKAROKI 360
Db 301 YFIYSQVEVYINFTDFASYEVVVDKPFLOCTRSIETGKTNNTCYTAGVCLLKAROKI 360
Qy 361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391
Db 361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391

RESULT 8
US-11-028-780-42
; Sequence 42, Application US/11028780
; Publication No. US20050163747A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: Novel Drosophila Tumor Necrosis Factor Class Molecule ("DmTNF")
; FILE REFERENCE: Variants Thereof
; FILE REFERENCE: D0016 DIV1
; CURRENT APPLICATION NUMBER: US/11/142,736
; CURRENT FILING DATE: 2005-06-01
; PRIOR FILING DATE: 2005-06-01
; PRIOR APPLICATION NUMBER: 60/190,816
; PRIOR FILING DATE: 2000-03-21
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-11-028-780-42

Query Match 100.0%; Score 2098; DB 6; Length 391;
```

```
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.,
; TITLE OF INVENTION: Heteromultimeric TNF Ligand Family members
; FILE REFERENCE: PF559C1
; CURRENT APPLICATION NUMBER: US/11/028,780
; CURRENT FILING DATE: 2005-01-05
; PRIOR APPLICATION NUMBER: 10/202,062
; PRIOR FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,838
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 42
; LENGTH: 391
; TYPE: PRT
; ORGANISM: human
US-11-028-780-42

Query Match 100.0%; Score 2098; DB 6; Length 391;
Best Local Similarity 100.0%; Pred. No. 2.6e-137;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGYPEVERRELLPAAAPRERGSGCGCGGAPARAGEGNSCLLFLGFFGLSLALHLLTLCC 60
Db 1 MGYPEVERRELLPAAAPRERGSGCGCGGAPARAGEGNSCLLFLGFFGLSLALHLLTLCC 60
Qy 61 YLELSELRRRGAESRLGGSGTPTGTSLSLGLDPSDPTITSHLGQSPKQPLEPGE 120
Db 61 YLELSELRRRGAESRLGGSGTPTGTSLSLGLDPSDPTITSHLGQSPKQPLEPGE 120
Qy 121 AALHSDSQDGHQMALNFFPDDEKPYSEESRRVRNRKRSKNEGADGVPVKNKKKKKAG 180
Db 121 AALHSDSQDGHQMALNFFPDDEKPYSEESRRVRNRKRSKNEGADGVPVKNKKKKKAG 180
Qy 181 PPGNGPPGPPGPPGPPGPPGPGIPIGPIGTVMGPPGPPGPPGPPGPPGPPGPPGPPG 240
Db 181 PPGNGPPGPPGPPGPPGPPGPGIPIGPIGTVMGPPGPPGPPGPPGPPGPPGPPGPPG 240
Qy 241 AGTRENQAVVHLQGGSAIQVKNLDSGGVLDNSRITMNPVKFKLHPRSGEVLVDGT 300
Db 241 AGTRENQAVVHLQGGSAIQVKNLDSGGVLDNSRITMNPVKFKLHPRSGEVLVDGT 300
Qy 301 YFIYSQVEVYINFTDFASYEVVVDKPFLOCTRSIETGKTNNTCYTAGVCLLKAROKI 360
Db 301 YFIYSQVEVYINFTDFASYEVVVDKPFLOCTRSIETGKTNNTCYTAGVCLLKAROKI 360
Qy 361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391
Db 361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391

RESULT 9
US-11-142-736-8
; Sequence 8, Application US/11142736
; Publication No. US20050227283A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: Novel Drosophila Tumor Necrosis Factor Class Molecule ("DmTNF")
; FILE REFERENCE: Variants Thereof
; FILE REFERENCE: D0016 DIV1
; CURRENT APPLICATION NUMBER: US/11/142,736
; CURRENT FILING DATE: 2005-06-01
; PRIOR FILING DATE: 2005-06-01
; PRIOR APPLICATION NUMBER: 60/190,816
; PRIOR FILING DATE: 2000-03-21
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-11-142-736-8

Query Match 100.0%; Score 2098; DB 6; Length 391;
```



Best Local Similarity 100.0%; Pred. No. 2.6e-137;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGYPEVERRELLPAAAPRRSGQCGCGAPARAGEGNSCLLFLGFFGLSLALHLLTLCC 60  
DB 1 MGYPEVERRELLPAAAPRRSGQCGCGAPARAGEGNSCLLFLGFFGLSLALHLLTLCC 60  
QY 61 YLELSELRRERGAESRLGSGTGTGTLSSGLGLDPDPSITSHLGQSPKQPLEPGE 120  
DB 61 YLELSELRRERGAESRLGSGTGTGTLSSGLGLDPDPSITSHLGQSPKQPLEPGE 120  
QY 121 AALHSDSQDHQWALLNFFPPDEKPYSEESRRVRNRKSKSNEGADGVPVKNKKGKAG 180  
DB 121 AALHSDSQDHQWALLNFFPPDEKPYSEESRRVRNRKSKSNEGADGVPVKNKKGKAG 180  
QY 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240  
DB 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240  
QY 241 AGTRENQPAVHLQGGSAIQVKNLDSGVLDWSRITMNPVKFKLHPRSGEVLVDGT 300  
DB 241 AGTRENQPAVHLQGGSAIQVKNLDSGVLDWSRITMNPVKFKLHPRSGEVLVDGT 300  
QY 301 YFIYSQVEYYINFTDFASYEVVVDKPFLOCTRSIETGKTNNTCYTAGVCLLKARQKI 360  
DB 301 YFIYSQVEYYINFTDFASYEVVVDKPFLOCTRSIETGKTNNTCYTAGVCLLKARQKI 360  
QY 361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391  
DB 361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391

RESULT 10

US-09-813-329-9  
; Sequence 9, Application US/09813329  
; Patent No. US20020012968A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: Novel Drosophila Tumor Necrosis Factor Class Mole  
; FILE OF INVENTION: Variants Thereof  
; FILE REFERENCE: D0016.np  
; CURRENT APPLICATION NUMBER: US/09/813,329  
; PRIOR FILING DATE: 2001-03-20  
; PRIOR APPLICATION NUMBER: 60/190,816  
; PRIOR FILING DATE: 2000-03-21  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 9  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-813-329-9

Query Match 93.7%; Score 1965; DB 3; Length 391;

Best Local Similarity 94.6%; Pred. No. 4.3e-128;  
Matches 370; Conservative 1; Mismatches 20; Indels 0; Gaps 0;

QY 1 MGYPEVERRELLPAAAPRRSGQCGCGAPARAGEGNSCLLFLGFFGLSLALHLLTLCC 60  
DB 1 MGYPEVERRELLPAAAPRRSGQCGCGAPARAGEGNSCLLFLGFFGLSLALHLLTLCC 60  
QY 61 YLELSELRRERGAESRLGSGTGTGTLSSGLGLDPDPSITSHLGQSPKQPLEPGE 120  
DB 61 YLELSELRRERGAESRLGSGTGTGTLSSGLGLDPDPSITSHLGQSPKQPLEPGE 120  
QY 121 AALHSDSQDHQWALLNFFPPDEKPYSEESRRVRNRKSKSNEGADGVPVKNKKGKAG 180  
DB 121 AALHSDSQDHQWALLNFFPPDEKPYSEESRRVRNRKSKSNEGADGVPVKNKKGKAG 180  
QY 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240  
DB 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240

QY 241 AGTRENQPAVHLQGGSAIQVKNLDSGVLDWSRITMNPVKFKLHPRSGEVLVDGT 300  
DB 241 AGTRENQPAVHLQGGSAIQVKNLDSGVLDWSRITMNPVKFKLHPRSGEVLVDGT 300  
QY 301 YFIYSQVEYYINFTDFASYEVVVDKPFLOCTRSIETGKTNNTCYTAGVCLLKARQKI 360  
DB 301 YFIYSQVEYYINFTDFASYEVVVDKPFLOCTRSIETGKTNNTCYTAGVCLLKARQKI 360  
QY 361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391  
DB 361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391

RESULT 11

US-11-142-736-9  
; Sequence 9, Application US/11142736  
; Publication No. US20050227283A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: Novel Drosophila Tumor Necrosis Factor Class Molecule ("DmTNP")  
; FILE OF INVENTION: Variants Thereof  
; FILE REFERENCE: D0016 DIV1  
; CURRENT APPLICATION NUMBER: US/11/142,736  
; PRIOR FILING DATE: 2005-06-01  
; PRIOR APPLICATION NUMBER: 60/190,816  
; PRIOR FILING DATE: 2000-03-21  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 9  
; LENGTH: 391  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-11-142-736-9

Query Match 93.7%; Score 1965; DB 6; Length 391;

Best Local Similarity 94.6%; Pred. No. 4.3e-128;  
Matches 370; Conservative 1; Mismatches 20; Indels 0; Gaps 0;

QY 1 MGYPEVERRELLPAAAPRRSGQCGCGAPARAGEGNSCLLFLGFFGLSLALHLLTLCC 60  
DB 1 MGYPEVERRELLPAAAPRRSGQCGCGAPARAGEGNSCLLFLGFFGLSLALHLLTLCC 60  
QY 61 YLELSELRRERGAESRLGSGTGTGTLSSGLGLDPDPSITSHLGQSPKQPLEPGE 120  
DB 61 YLELSELRRERGAESRLGSGTGTGTLSSGLGLDPDPSITSHLGQSPKQPLEPGE 120  
QY 121 AALHSDSQDHQWALLNFFPPDEKPYSEESRRVRNRKSKSNEGADGVPVKNKKGKAG 180  
DB 121 AALHSDSQDHQWALLNFFPPDEKPYSEESRRVRNRKSKSNEGADGVPVKNKKGKAG 180  
QY 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240  
DB 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240  
QY 241 AGTRENQPAVHLQGGSAIQVKNLDSGVLDWSRITMNPVKFKLHPRSGEVLVDGT 300  
DB 241 AGTRENQPAVHLQGGSAIQVKNLDSGVLDWSRITMNPVKFKLHPRSGEVLVDGT 300  
QY 301 YFIYSQVEYYINFTDFASYEVVVDKPFLOCTRSIETGKTNNTCYTAGVCLLKARQKI 360  
DB 301 YFIYSQVEYYINFTDFASYEVVVDKPFLOCTRSIETGKTNNTCYTAGVCLLKARQKI 360  
QY 361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391  
DB 361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391

RESULT 12

US-09-729-658B-4  
; Sequence 4, Application US/09729658B  
; Publication No. US20030023991A1  
; GENERAL INFORMATION:  
; APPLICANT: Zonana et al.

; TITLE OF INVENTION: Hypohydrotic ectodermal dysplasia genes and proteins
; FILE REFERENCE: 55924
; CURRENT APPLICATION NUMBER: US/09/729,658B
; CURRENT FILING DATE: 2000-12-04
; PRIOR FILING DATE: 09/342,681
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 60/092,279
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: 60/112,366
; PRIOR FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-729-658B-4

Query Match 89.3%; Score 1873; DB 3; Length 377;
Best Local Similarity 90.8%; Pred. No. 1e-121;
Matches 355; Conservative 2; Mismatches 20; Indels 14; Gaps 1;

; 1 MGYEVEVERRELLPAAAPRRSGSQCGCGAPARAGEGNSCLLPLGFFGLSLALHLLTLC 60
; 1 MGYEVEVERREPLPAAAPRRSGSQCGCGAPARAGEGNSCLLPLGFFGLSLALHLLTLC 60
; 61 YLEIRSELRRRGAESRLGSGTGTGTSLSLGLDPSITSHLGOPSPKQPLEPGE 120
; 61 YLEIRSELRRRGTESRLGSGAPGTGTSLSLGLDPSITSHLGOPSPKQPLEPGE 120
; 121 AALHSDSDQHOMALLNFFPFDEKPYSEERVRNRKRSKNEGADGPVKNKKKGKAG 180
; 121 DPLPPESQDRHOMALLNFFPFDEKAYSEERVRNRKRSKNEGADGPVKNKKKGKAG 180
; 181 PPGNGPPGPPGPPGPPGPGIPGIPGTTVMGPPGPPGPPGPPGPPGPPGPPGPPG 240
; 181 PPGNGPPGPPGPPGPPGPPGIPGIPGTTVMGPPGPPGPPGPPGPPGPPGPPGPPG 240
; 241 AGTRENQAVVHLQOGSAIQVKNLDSGVLDNSRITMNPVKVFKLHPRSGEVLVDGT 300
; 241 TGTRENQAVVHLQOGSAIQVKNLDSGVLDNSRITMNPVKVFKLHPRSGEL----- 293
; 301 YPIYSQVEVYINFTDFASYEVVVVDEKPFLOCTRSIETGKTNNTCYTAGVCLLKARQKI 360
; 294 -----EYVYINFTDFASYEVVVVDEKPFLOCTRSIETGKTNNTCYTAGVCLLKARQKI 346
; 361 AVKMVHADISINMSKHTTFFGAIIRLGEAPAS 391
; 347 AVKMVHADISINMSKHTTFFGAIIRLGEAPAS 377

RESULT 13
US-10-503-999-6
; Sequence 6, Application US/10503999
; Publication No. US2005015287A1
; GENERAL INFORMATION:
; APPLICANT: Gaide, Olivier
; APPLICANT: Schneider, Pascal
; APPLICANT: Tschopp, Jurg
; TITLE OF INVENTION: Fusion Constructs Containing Active Sections Of TNF Ligands
; FILE REFERENCE: 11436\*13
; CURRENT APPLICATION NUMBER: US/10/503,999
; CURRENT FILING DATE: 2004-08-09
; PRIOR APPLICATION NUMBER: PCT/EP02/09354
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: DE 102 05 368.5
; PRIOR FILING DATE: 2002-02-10
; PRIOR APPLICATION NUMBER: DE 102 05 583.1
; PRIOR FILING DATE: 2002-02-11
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 410
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
; US-10-503-999-8

Query Match 36.2%; Score 759; DB 5; Length 410;
Best Local Similarity 86.4%; Pred. No. 1.9e-44;
Matches 153; Conservative 1; Mismatches 15; Indels 8; Gaps 3;

; 215 GPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 274
; 242 GKRSPPQPKPQPKP---EPEGSLQ---VDENQPAVVHLQOGSAIQVKNLDSGVLDNW 295
; 275 SRITMNPVKVFKLHPRSGEVLVDGTGYFIYSQVEVYINFTDFASYEVVVVDEKPFLOCTR 334
; 296 SRITMNPVKVFKLHPRSGEVLVDGTGYFIYSQ--VYINFTDFASYEVVVVDEKPFLOCTR 353
; 335 SIETGKTNNTCYTAGVCLLKARQKIAVKMVHADISINMSKHTTFFGAIIRLGEAPAS 391
; 354 SIETGKTNNTCYTAGVCLLKARQKIAVKMVHADISINMSKHTTFFGAIIRLGEAPAS 410

RESULT 15
US-10-185-425-9
; Sequence 9, Application US/10185425
; Publication No. US20030053984A1
; GENERAL INFORMATION:

; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
; US-10-503-999-6

Query Match 37.1%; Score 779; DB 5; Length 412;
Best Local Similarity 87.6%; Pred. No. 7.9e-46;
Matches 155; Conservative 1; Mismatches 15; Indels 6; Gaps 2;

; 215 GPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 274
; 242 GKRSPPQPKPQPKP---EPEGSLQ---VDENQPAVVHLQOGSAIQVKNLDSGVLDNW 295
; 275 SRITMNPVKVFKLHPRSGEVLVDGTGYFIYSQVEVYINFTDFASYEVVVVDEKPFLOCTR 334
; 296 SRITMNPVKVFKLHPRSGEVLVDGTGYFIYSQVEVYINFTDFASYEVVVVDEKPFLOCTR 355
; 335 SIETGKTNNTCYTAGVCLLKARQKIAVKMVHADISINMSKHTTFFGAIIRLGEAPAS 391
; 356 SIETGKTNNTCYTAGVCLLKARQKIAVKMVHADISINMSKHTTFFGAIIRLGEAPAS 412

RESULT 14
US-10-503-999-8
; Sequence 8, Application US/10503999
; Publication No. US2005015287A1
; GENERAL INFORMATION:
; APPLICANT: Gaide, Olivier
; APPLICANT: Schneider, Pascal
; APPLICANT: Tschopp, Jurg
; TITLE OF INVENTION: Fusion Constructs Containing Active Sections Of TNF Ligands
; FILE REFERENCE: 11436\*13
; CURRENT APPLICATION NUMBER: US/10/503,999
; CURRENT FILING DATE: 2004-08-09
; PRIOR APPLICATION NUMBER: PCT/EP02/09354
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: DE 102 05 368.5
; PRIOR FILING DATE: 2002-02-10
; PRIOR APPLICATION NUMBER: DE 102 05 583.1
; PRIOR FILING DATE: 2002-02-11
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 410
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
; US-10-503-999-8

Query Match 36.2%; Score 759; DB 5; Length 410;
Best Local Similarity 86.4%; Pred. No. 1.9e-44;
Matches 153; Conservative 1; Mismatches 15; Indels 8; Gaps 3;

; 215 GPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 274
; 242 GKRSPPQPKPQPKP---EPEGSLQ---VDENQPAVVHLQOGSAIQVKNLDSGVLDNW 295
; 275 SRITMNPVKVFKLHPRSGEVLVDGTGYFIYSQVEVYINFTDFASYEVVVVDEKPFLOCTR 334
; 296 SRITMNPVKVFKLHPRSGEVLVDGTGYFIYSQ--VYINFTDFASYEVVVVDEKPFLOCTR 353
; 335 SIETGKTNNTCYTAGVCLLKARQKIAVKMVHADISINMSKHTTFFGAIIRLGEAPAS 391
; 354 SIETGKTNNTCYTAGVCLLKARQKIAVKMVHADISINMSKHTTFFGAIIRLGEAPAS 410

RESULT 15
US-10-185-425-9
; Sequence 9, Application US/10185425
; Publication No. US20030053984A1
; GENERAL INFORMATION:



***This Page Blank (uspto)***

```
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 912
; LENGTH: 1806
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-912

Query Match      12.1%; Score 254; DB 6; Length 1806;
Best Local Similarity 32.0%; Pred. No. 6.2e-12;
Matches 93; Conservative 15; Mismatches 87; Indels 96; Gaps 17;

Qy 13 PAAAPRGRSGQC-----GCCGAPARAGEGNSCLLFLGFFGLSLALHLLTLCCLYLRSEL 68
Db 755 PIGYVPRGVKGADGVRLKSGKEGEDG----FPGFKG-DMGLK-----GDRGEV 801
Qy 69 -----RRERGAE---SRLGSGTPTGTSGLSSLGGLD-PDSPITSHLQSPKQOPLPG 119
Db 802 GQIGPRGEDGPEGPKRAGTGDPRGSGQAGEKGLGVPLP--GYFGRQPKGSTGTFPG 859
Qy 120 EAAHSDSQDGHQWALLNFFPFDEKPYSEESRRVR-----RNKRS-----KSNEGADGPV 170
Db 860 -----FPGAN--GEKGARGVAGKPGPRGQGTGPRGSRGARGPT 897
Qy 171 -KNKKKGKKA-----GPPGPNGGPPGPPG---PGQPPGPIPGIPG-----TTVM 214
Db 898 GKPGPKTSGDGGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 957
Qy 215 GPPGPPGP-----PGQPPGLQGPSGAADKAGTR 244
Db 958 KGTGPPGPGVGVGPGQGTGTGPIGRGHGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 1008

RESULT 14
US-10-995-561-915
; Sequence 915, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 915
; LENGTH: 1806
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-915

Query Match      12.1%; Score 254; DB 6; Length 1806;
Best Local Similarity 32.0%; Pred. No. 6.2e-12;
Matches 93; Conservative 15; Mismatches 87; Indels 96; Gaps 17;

Qy 13 PAAAPRGRSGQC-----GCCGAPARAGEGNSCLLFLGFFGLSLALHLLTLCCLYLRSEL 68
Db 755 PIGYVPRGVKGADGVRLKSGKEGEDG----FPGFKG-DMGLK-----GDRGEV 801
Qy 69 -----RRERGAE---SRLGSGTPTGTSGLSSLGGLD-PDSPITSHLQSPKQOPLPG 119
Db 802 GQIGPRGEDGPEGPKRAGTGDPRGSGQAGEKGLGVPLP--GYFGRQPKGSTGTFPG 859
Qy 120 EAAHSDSQDGHQWALLNFFPFDEKPYSEESRRVR-----RNKRS-----KSNEGADGPV 170
Db 860 -----FPGAN--GEKGARGVAGKPGPRGQGTGPRGSRGARGPT 897

RESULT 15
US-10-995-561-910
; Sequence 910, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 910
; LENGTH: 1818
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-910

Query Match      12.1%; Score 254; DB 6; Length 1818;
Best Local Similarity 32.0%; Pred. No. 6.2e-12;
Matches 93; Conservative 15; Mismatches 87; Indels 96; Gaps 17;

Qy 13 PAAAPRGRSGQC-----GCCGAPARAGEGNSCLLFLGFFGLSLALHLLTLCCLYLRSEL 68
Db 767 PIGYVPRGVKGADGVRLKSGKEGEDG----FPGFKG-DMGLK-----GDRGEV 813
Qy 69 -----RRERGAE---SRLGSGTPTGTSGLSSLGGLD-PDSPITSHLQSPKQOPLPG 119
Db 814 GQIGPRGEDGPEGPKRAGTGDPRGSGQAGEKGLGVPLP--GYFGRQPKGSTGTFPG 871
Qy 120 EAAHSDSQDGHQWALLNFFPFDEKPYSEESRRVR-----RNKRS-----KSNEGADGPV 170
Db 872 -----FPGAN--GEKGARGVAGKPGPRGQGTGPRGSRGARGPT 909
Qy 171 -KNKKKGKKA-----GPPGPNGGPPGPPG---PGQPPGPIPGIPG-----TTVM 214
Db 910 KTFGPKTSGDGGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 969
Qy 215 GPPGPPGP-----PGQPPGLQGPSGAADKAGTR 244
Db 970 KGTGPPGPGVGVGPGQGTGTGPIGRGHGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 1020

Search completed: January 13, 2006, 16:26:34
Job time : 32 secs
```

; APPLICANT: BURGART, LAWRENCE J.		
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND		
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND		
; TITLE OF INVENTION: THERAPY OF COLON CANCER		
; FILE REFERENCE: MP01-029P2RNM		
; CURRENT APPLICATION NUMBER: US/11/186,284		
; CURRENT FILING DATE: 2005-07-21		
; PRIOR APPLICATION NUMBER: US/10/301,822		
; PRIOR FILING DATE: 2002-11-21		
; PRIOR APPLICATION NUMBER: US 60/339,971		
; PRIOR FILING DATE: 2001-12-10		
; PRIOR APPLICATION NUMBER: US 60/361,978		
; PRIOR FILING DATE: 2002-03-05		
; PRIOR APPLICATION NUMBER: US 60/381,988		
; PRIOR FILING DATE: 2002-05-20		
; NUMBER OF SEQ ID NOS: 228		
; SOFTWARE: FASTSEQ for Windows Version 4.0		
; SEQ ID NO 33		
; LENGTH: 1466		
; TYPE: PRT		
; ORGANISM: Homo sapiens		
US-11-186-284-33		
Query Match 12.2%; Score 256; DB 7; Length 1466;		
Best Local Similarity 28.8%; Pred. No. 3.4e-12;		
Matches 86; Conservative 12; Mismatches 97; Indels 104; Gaps 10;		
Qy	13	PAAPRERGSGQ-----CCGGAPARAGGNSCLLFLGFFGLSLALHLLTLCCLYLRSEL 68
Db	793	PRGSGERGETGPPGAGFPAGQNGRPGG-----823
Qy	69	RRERGAESRLGSGTPTGTSGLTSSLGGLDPSPTITSHLQSPSPKQOPLPGEAALHSDSQ 128
Db	824	KGERGAPGKEGGPPGVAGPPGSGPAGPPGP-----QGVKGERSGPGPGAAG-----873
Qy	129	DGHQALLNFPFDEKPYSEESRRVRNKSNEGADGPVNKKKXKAGPPGPNPPP 188
Db	874	-----FPG-----ARGLPGPGSGNGNPGPPSPGSGPGKDGPPGAGNTGAP 914
Qy	189	GPP-----GPPGPPGPPGIPGIPGTT-----VMGPPGPPGP-----PGPQ 226
Db	915	GSPGVGPKGDAGQPGKEKSPQAQPPGAPGLGAGTGTAGRLAGPFGMFGPRGSPGPQ 974
Qy	227	-----GPPQLQSPGAADKAGT--RENQPAVVHLQOQSAIVQKND 265
Db	975	GVKSGSKPGANGLSGRGGPPGQGLPLGLAGTAGPRGDRGNPGSDGLPCRDSGPGKGD 1033
RESULT 11		
US-10-995-561-911		
; Sequence 911, Application US/10995561		
; Publication No. US20050272054A1		
; GENERAL INFORMATION:		
; APPLICANT: CARGILL, Michele et al.		
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH		
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF		
; TITLE OF INVENTION: DETECTION AND USES THEREOF		
; FILE REFERENCE: CL001559		
; CURRENT APPLICATION NUMBER: US/10/995,561		
; CURRENT FILING DATE: 2004-11-24		
; NUMBER OF SEQ ID NOS: 85702		
; SOFTWARE: FastSeq for Windows Version 4.0		
; SEQ ID NO 911		
; LENGTH: 1767		
; TYPE: PRT		
; ORGANISM: Homo sapiens		
US-10-995-561-911		
Query Match 12.1%; Score 254; DB 6; Length 1767;		
Best Local Similarity 32.0%; Pred. No. 6e-12;		
Matches 93; Conservative 15; Mismatches 87; Indels 96; Gaps 17;		
Qy	13	PAAPRERGSGC-----CCGGAPARAGGNSCLLFLGFFGLSLALHLLTLCCLYLRSEL 68
RESULT 13		
US-10-995-561-912		
; Sequence 912, Application US/10995561		
; Publication No. US20050272054A1		
; GENERAL INFORMATION:		
; APPLICANT: CARGILL, Michele et al.		
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH		
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF		
; TITLE OF INVENTION: DETECTION AND USES THEREOF		
; FILE REFERENCE: CL001559		
; CURRENT APPLICATION NUMBER: US/10/995,561		
; CURRENT FILING DATE: 2004-11-24		
; NUMBER OF SEQ ID NOS: 85702		
; SOFTWARE: FastSeq for Windows Version 4.0		
; SEQ ID NO 911		
; LENGTH: 1767		
; TYPE: PRT		
; ORGANISM: Homo sapiens		
US-10-995-561-911		
Query Match 12.1%; Score 254; DB 6; Length 1767;		
Best Local Similarity 32.0%; Pred. No. 6e-12;		
Matches 93; Conservative 15; Mismatches 87; Indels 96; Gaps 17;		
Qy	13	PAAPRERGSGC-----CCGGAPARAGGNSCLLFLGFFGLSLALHLLTLCCLYLRSEL 68



```

; TYPE: PRT
; ORGANISM: Homo sapiens
; US-11-186-284-28

Query Match      12.7%; Score 265.5; DB 7; Length 1464;
Best Local Similarity 29.7%; Pred. No. 6.3e-13;
Matches 78; Conservative 16; Mismatches 80; Indels 89; Gaps 9;

Qy 13 PAAAPRERSGQ-----CGCGGAPARAGE-----GNSCLLFLGFFGLSLALHLTLCCY 61
Db 774 PAGAPGDKGESGSPAGPTGARGAPCDRGEPGPGA-----GFAG-----815

Qy 62 LELSELRRERGAESRLGGSGTGTGTLSSLGGLDPSITSHLGQSPKQOPLPGEA 121
Db 816 -----PPGADGQFGAKGEFDAGAKDAGPPGAPG-----AGPPGPIGNVAGPGA- 861

Qy 122 ALHSDSDQHQMALLNFFPDEKPYSEESRRVRRNRKSKSNEGADGVPVKNKKKKKAGP 181
Db 862 -----KGARGSAGPPGATGFPGAAGRVP 885

Qy 182 PGNGPPGPPGPPG-----QPPGPIGPIGPIGTTVMGPPGPPGPPGPPGPPGPPG 235
Db 886 PGPSGNAGPPGPPGPKAGKEGKGPRGTGPAGRPGE--VGPPGPPGPKAGEKSGFGADGPA 943

Qy 236 GAADKAGTR--ENQPAVVHLQO 256
Db 944 GAGTTPGQGIAGQGVVGLPGQ 966

RESULT 5
US-11-021-603-2
; Sequence 2, Application US/11021603
; Publication No. US20060003954A1
; GENERAL INFORMATION:
; APPLICANT: Beri, Rajinder
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES
; FILE REFERENCE: 06275-254U51
; CURRENT APPLICATION NUMBER: US/11/021,603
; CURRENT FILING DATE: 2004-12-21
; PRIOR FILING DATE: US/10/149,352
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: PCT/GB00/04741
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: GB 9929487.8
; PRIOR FILING DATE: 1999-12-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 4.0
; SEQ ID NO 2
; LENGTH: 1464
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-021-603-2

Query Match      12.7%; Score 265.5; DB 7; Length 1464;
Best Local Similarity 29.7%; Pred. No. 6.3e-13;
Matches 78; Conservative 16; Mismatches 80; Indels 89; Gaps 9;

Qy 13 PAAAPRERSGQ-----CGCGGAPARAGE-----GNSCLLFLGFFGLSLALHLTLCCY 61
Db 774 PAGAPGDKGESGSPAGPTGARGAPCDRGEPGPGA-----GFAG-----815

Qy 62 LELSELRRERGAESRLGGSGTGTGTLSSLGGLDPSITSHLGQSPKQOPLPGEA 121
Db 816 -----PPGADGQFGAKGEFDAGAKDAGPPGAPG-----AGPPGPIGNVAGPGA- 861

Qy 122 ALHSDSDQHQMALLNFFPDEKPYSEESRRVRRNRKSKSNEGADGVPVKNKKKKKAGP 181
Db 862 -----KGARGSAGPPGATGFPGAAGRVP 885

Qy 182 PGNGPPGPPGPPG-----QPPGPIGPIGPIGTTVMGPPGPPGPPGPPGPPG 235
Db 886 PGPSGNAGPPGPPGPKAGKEGKGPRGTGPAGRPGE--VGPPGPPGPKAGEKSGFGADGPA 943

Qy 236 GAADKAGTR--ENQPAVVHLQO 256
Db 944 GAGTTPGQGIAGQGVVGLPGQ 966

RESULT 5
US-11-021-603-2
; Sequence 2, Application US/11021603
; Publication No. US20060003954A1
; GENERAL INFORMATION:
; APPLICANT: Beri, Rajinder
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES
; FILE REFERENCE: 06275-254U51
; CURRENT APPLICATION NUMBER: US/11/021,603
; CURRENT FILING DATE: 2004-12-21
; PRIOR FILING DATE: US/10/149,352
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: PCT/GB00/04741
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: GB 9929487.8
; PRIOR FILING DATE: 1999-12-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 4.0
; SEQ ID NO 2
; LENGTH: 1464
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-021-603-2

Query Match      12.7%; Score 265.5; DB 7; Length 1464;
Best Local Similarity 29.7%; Pred. No. 6.3e-13;
Matches 78; Conservative 16; Mismatches 80; Indels 89; Gaps 9;

Qy 13 PAAAPRERSGQ-----CGCGGAPARAGE-----GNSCLLFLGFFGLSLALHLTLCCY 61
Db 774 PAGAPGDKGESGSPAGPTGARGAPCDRGEPGPGA-----GFAG-----815

Qy 62 LELSELRRERGAESRLGGSGTGTGTLSSLGGLDPSITSHLGQSPKQOPLPGEA 121
Db 816 -----PPGADGQFGAKGEFDAGAKDAGPPGAPG-----AGPPGPIGNVAGPGA- 861

Qy 122 ALHSDSDQHQMALLNFFPDEKPYSEESRRVRRNRKSKSNEGADGVPVKNKKKKKAGP 181
Db 862 -----KGARGSAGPPGATGFPGAAGRVP 885

Qy 182 PGNGPPGPPGPPG-----QPPGPIGPIGPIGTTVMGPPGPPGPPGPPGPPG 235
Db 886 PGPSGNAGPPGPPGPKAGKEGKGPRGTGPAGRPGE--VGPPGPPGPKAGEKSGFGADGPA 943

```



```
RESULT 2
US-10-821-234-914
; Sequence 914, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 914
; LENGTH: 1532
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-914

Query Match 13.0%; Score 273; DB 6; Length 1532;
Best Local Similarity 31.2%; Pred. No. 1.8e-13;
Matches 85; Conservative 13; Mismatches 64; Indels 110; Gaps 13;

Qy 20 RGSQC-----CGGAPARAGNSCLLPFGFGLSLALHLLTCLCYELRSELRRERGA 75
Db 789 RGEQGLTGMPGIRPPGSGDP-----GKPGLT-----GPK 819

Qy 76 SRLGSGTPTGTSGLSLGLDPSPTSH-----LQSPKQOPLPG 119
Db 820 GPQGLGTPGPRGKIGPGA--PGKIVTSESSMLTVPGPPGCGAMGPPGPGAGPAG 877

Qy 120 EALHSDSDQGHQMALNFFFPDEKPYSEESRRVRNRKSKNEGADGPKVKKKKGKA 179
Db 878 PAGL-----PGHQEVL-----NLQGP-----P 894

Qy 180 GPPGNGPPGP--PGPPGQPP--GIPGIPG-----TTVMGPPGPPGPPGP--- 225
Db 895 GPPGPRGPPGSIPTGPPGPRGPGGLPGPPGPGSFLNSSETFLFGPPGPPGPKGD 954

Qy 226 QGPPGLQPSGAADKAG--TRENQPAVHLQ 255
Db 955 QGPPGPRGHQGEQGLPGFSTGSSGSFGLNLQ 986

RESULT 3
US-11-000-463-243
; Sequence 243, Application US/11000463
; Publication No. US20050266423A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Chen, Rui-hong
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Zhou, Ping
; APPLICANT: Cao, Yi-Cheng
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 785CIP4N
; CURRENT APPLICATION NUMBER: US/11/000,463
; CURRENT FILING DATE: 2004-11-29
; PRIOR APPLICATION NUMBER: 10/291,265
; PRIOR FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/922,279
```

```
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 243
; LENGTH: 1464
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-000-463-243

Query Match 12.7%; Score 265.5; DB 7; Length 1464;
Best Local Similarity 29.7%; Pred. No. 6.3e-13;
Matches 78; Conservative 16; Mismatches 80; Indels 89; Gaps 9;

Qy 13 PAAAPRRGSGQ-----CGCGAPARAGE-----GNSCLLPFGFGLSLALHLLTLCY 61
Db 774 PAGAPGDKGESGSPGAPGTCARGAGDGRGPPGPA-----GFAG----- 815

Qy 62 LELRSELRRERGAESRLGSGTPTGTSGLGLDPSPTSHLGLGSPSPKQPLEPGEA 121
Db 816 -----PPGADGQFGAKGEPDAGAKGDAGPPGAGP-----AGPPGPIGNVGAFA- 861

Qy 122 ALHSDSDQGHQMALNFFFPDEKPYSEESRRVRNRKSKNEGADGPKVKKKKGKA 181
Db 862 -----KGARSGAGPPGATGFPGAAGRVP 885

Qy 182 PGNPGPPGPPGPPG-----QGPPGIPGIPGIPGTTVMGPPGPPGPPGPPGPPG 235
Db 886 PGPSGNAGPPGPPGPPGPKGKGKPRGTGPAGRFE--VGPPGPPGPAKGKSGADGPA 943

Qy 236 GAADKAGTR--ENQPAVVHLQ 256
Db 944 GAPGTPGQGIAGQGVVGLPG 966

RESULT 4
US-11-186-284-28
; Sequence 28, Application US/11186284
; Publication No. US20050266493A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: BURGART, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MPM01-029P2RNM
; CURRENT APPLICATION NUMBER: US/11/186,284
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 1464
```

